Computational Economics
https://doi.org/10.1007/s10614-021-10200-y

DSGE-SVt: An Econometric Toolkit for High-Dimensional DSGE Models with SV and *t* Errors

Siddhartha Chib[1] · **Minchul Shin**[2] · **Fei Tan**[3,4]

Accepted: 10 September 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Presently there is growing interest in dynamic stochastic general equilibrium (DSGE) models with more parameters, endogenous variables, exogenous shocks, and observable variables than the Smets and Wouters (Am Econ Rev 97(3):586–606, 2007) model, and the incorporation of non-Gaussian distribution and time-varying volatility. A primary goal of this paper is to introduce a user-friendly MATLAB toolkit designed to reliably estimate such high-dimensional models. It simulates the posterior distribution by the tailored random block Metropolis-Hastings (TaRB-MH) algorithm of Chib and Ramamurthy (J Econom 155(1):19–38, 2010), calculates the marginal likelihood by the method of Chib (J Am Stat Assoc 90:1313–1312, 1995) and Chib and Jeliazkov (J Am Stat Assoc 96(453):270–281, 2001), and includes various post-estimation tools that are important for policy analysis, for example, functions for generating point and density forecasts. We also introduce two novel features, i.e., tailoring-at-random-frequency and parallel computing, to boost the overall computational efficiency. Another goal is to provide pointers on the prior, estimation, and comparison of these DSGE models. To demonstrate the performance of our toolkit, we apply it to estimate an extended version of the new Keynesian model of Leeper et al (Am Econom Rev 107(8):2409–2454, 2017) that has 51 parameters, 21 endogenous variables, 8 exogenous shocks, 8 observable variables, and 1494 non-Gaussian and nonlinear latent variables.

Keywords DSGE models · Bayesian inference · Marginal likelihood · Tailored proposal densities · Random blocks · Student-*t* shocks · Stochastic volatility

JEL Classification C11 · C15 · C32 · E37 · E63

The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System.

Extended author information available on the last page of the article

Published online: 13 October 2021

1 Introduction

Over the past 20 years or so, dynamic stochastic general equilibrium (DSGE) models have become the mainstay of macroeconomic policy analysis and forecasting. Presently there is growing interest in DSGE models with more parameters, endogenous variables, exogenous shocks, and observable variables than the Smets and Wouters (2007) model, and the incorporation of non-Gaussian distribution, as in Chib and Ramamurthy (2014) and Cúrdia et al. (2014), and time-varying volatility, as in Justiniano and Primiceri (2008).[1] This is because these higher-dimensional DSGE models are more realistic and have the potential to provide better statistical fit to the data. Despite wide spread use of Bayesian estimation techniques, based on Markov chain Monte Carlo (MCMC) simulation methods [see Chib and Greenberg (1995b) and Herbst and Schorfheide (2016) for further details about these methods], the estimation of high-dimensional DSGE models is challenging. The popular DYNARE software, which has proved useful for small and medium-scale models is, however, currently not capable of handling the preceding DSGE models, thus inhibiting the formulation, estimation, and comparison of such models for policy analysis and prediction.

A primary goal of this paper is to introduce a user-friendly MATLAB toolkit for estimating high-dimensional DSGE models with stochastic volatility (SV) and Student-t shocks. Outside of the DSGE context, SV-Student-t models were first studied by Chib et al. (2002) and applied to financial data, then by Cúrdia et al. (2014) with applications to macroeconomic time series. Incorporating SV-Student-t shocks in DSGE models, however, introduces new challenges because of the complex mapping from the structural parameters to those of the state space model that emerges from the rational expectations solution of the equilibrium conditions. Our toolkit relies on the tailored random block Metropolis-Hastings (TaRB-MH) algorithm of Chib and Ramamurthy (2010) to deal with these challenges. Recent applications of this algorithm to DSGE models include, e.g., Born and Pfeifer (2014), Rathke et al. (2017), Kulish et al. (2017), and Kapetanios et al. (2019), while applications to other problems in economics include Kim and Kang (2019) and Mele (2020), amongst many others. We also introduce the tailoring-at-random-frequency procedure to accelerate the posterior sampling without losing the validity and efficiency of the original algorithm. The TaRB-MH algorithm may appear to require work, but random blocking and tailoring are central to generating efficient exploration of the posterior distribution. The TaRB-MH algorithm is also available in DYNARE, but only for models without Student-t shocks and stochastic volatility. Even there, we have found in experiments that its implementation is not as efficient as the one in our toolkit.

The marginal likelihood (the integral of the sampling density over the prior of the parameters) plays a central role in Bayesian model comparisons. In our toolkit we calculate this quantity by a parallel implementation of the Chib (1995) and Chib and Jeliazkov (2001) method. The marginal likelihood is also available in DYNARE, but

[1] See also, e.g., Dave and Malik (2017), Chiu et al. (2017), Franta (2017), and Liu (2019) for macroeconomic implications of fat-tailed shocks and stochastic volatility.

it is obtained by a modified version of the Gelfand and Dey (1994) method [also see, e.g., Justiniano and Primiceri (2008) and Cúrdia et al. (2014) for use of this method in DSGE models with Student-*t* shocks and stochastic volatility]. The latter method, however, is not as reliable as the Chib and Jeliazkov (2001) method. It is subject to upward finite-sample bias in models with latent variables and runs the risk of misleading model comparisons [see Sims et al. (2008) and Chan and Grant (2015) for such examples]. As this point is not well recognized in the DSGE model literature, we document its performance in simulated examples. It is shown to mistakenly favor models with fatter tails and incorrect time-varying variance dynamics. Finally, our toolkit includes various post-estimation tools that are important for policy analysis, for example, functions for generating point and density forecasts.

Another goal is to provide pointers on dealing with high-dimensional DSGE models that promote more reliable estimation and that are incorporated by default in our toolkit. Due to the complex mapping from the structural parameters to those of the state space form, standard prior assumptions about structural parameters may still imply a distribution of the data that is strongly at odds with actual observations. To see if this is the case, we sample the prior many times, solve for the equilibrium solution, and then sample the endogenous variables. Second, we suggest the use of a training sample to fix the hyperparameters. Although training sample priors are common in the vector autoregression (VAR) literature, they are not typically used in the DSGE setting. We also suggest the use of the Student-*t* family of distributions as the prior family for the location parameters. This tends to further mitigate the possibility of prior-sample conflicts and leads to more robust results. Finally, we invest in the most efficient way of sampling different blocks, for example, sampling the non-structural parameters and the latent variables by the integration sampler of Kim et al. (1998).

The rest of the paper is organized as follows. The next section outlines a prototypical high-dimensional DSGE model for the subsequent analysis. Sections 3–5 provide pointers on prior construction, posterior sampling, and model comparison accompanied by both empirical results and simulation evidence. Section 6 conducts an out-of-sample forecast analysis. Section 7 concludes. The appendix contains a detailed summary of the high-dimensional DSGE model (Appendix "A"), a small-scale DSGE model used in Sect. 6 (Appendix "B"), and a practical user guide on how to run our MATLAB toolkit called 'DSGE-SVt' (Appendix "C").[2]

2 High-Dimensional DSGE Model

As a template, we consider an extended version of the new Keynesian model of Leeper et al. (2017) that includes both fat-tailed shocks and time-varying volatility. This high-dimensional DSGE model consists of 51 parameters, 21 endogenous variables, 8 exogenous shocks, 8 observable variables, and 1, 494 non-Gaussian and nonlinear latent variables. For model comparison purposes, we also follow Leeper et al. (2017) and consider two distinct regimes of the policy parameter space—regime-M and regime-F—that deliver unique bounded rational expectations equilibria.

[2] The toolkit is publicly available at https://github.com/econdojo/dsge-svt. The results reported in this paper can be replicated by running the program `demo.m`.

The Leeper-Traum-Walker model consists of 36 log-linearized equilibrium equations and can be expressed generically in the form

$$\underset{(36\times36)}{\Gamma_0(\theta^S)}x_t = \underset{(36\times36)}{\Gamma_1(\theta^S)}x_{t-1} + \underset{(36\times8)}{\Psi}\epsilon_t + \underset{(36\times7)}{\Pi}\eta_t \tag{2.1}$$

where θ^S is a vector of 27 structural parameters, x_t is a vector of 36 model variables, ϵ_t is a vector of 8 shock innovations, η_t is a vector of 7 forecast errors, and $(\Gamma_0, \Gamma_1, \Psi, \Pi)$ are coefficient matrices with dimensions indicated below them.

Suppose now, following Chib and Ramamurthy (2014), that the shock innovations follow a multivariate Student-t distribution, i.e., $\epsilon_t \sim t_v(0, \Sigma_t)$. Here v denotes the degrees of freedom and Σ_t is a diagonal matrix with the time-varying volatility $\sigma^2_{s,t}$ for each individual innovation ϵ^s_t on its main diagonal, where s refers to the shock index.[3] For estimation convenience, it is useful to represent each element of ϵ_t as a mixture of normals by introducing a Gamma distributed random variable λ_t,

$$\epsilon^s_t = \lambda_t^{-1/2} e^{h^s_t/2}\varepsilon^s_t, \qquad \lambda_t \sim \mathbb{G}\left(\frac{v}{2}, \frac{v}{2}\right), \qquad \varepsilon^s_t \sim \mathbb{N}(0,1) \tag{2.2}$$

For further realism, following Kim et al. (1998), the logarithm of each volatility $h^s_t = \ln \sigma^2_{s,t}$ collected in an 8×1 vector h_t evolves as a stationary ($|\phi_s| < 1$) process

$$h^s_t = (1-\phi_s)\mu_s + \phi_s h^s_{t-1} + \eta^s_t, \qquad \eta^s_t \sim \mathbb{N}(0, \omega^2_s) \tag{2.3}$$

where we collect all the volatility parameters $(\mu_s, \phi_s, \omega^2_s)$ in a 24×1 vector θ^V.

Equations (2.1)–(2.3) is the type of high-dimensional DSGE model that is of much contemporary interest. This model is completed with a measurement equation that connects the state variables x_t to the observable measurements y_t. Then, given the sample data $y_{1:T}$ on an 8×1 vector y_t for periods $t = 1, \ldots, T$, the goal is to learn about (i) the model parameters $\theta = (\theta^S, \theta^V)$, (ii) the non-Gaussian latent variables $\lambda_{1:T}$ needed for modeling Student-t shocks, and (iii) the nonlinear latent variables $h_{1:T}$ representing log volatilities. By now, the general framework for doing this inference, based primarily on Bayesian tools, is quite well established. If $\pi(\theta)$ denotes the prior distribution, then MCMC methods are used to provide sample draws of the augmented posterior distribution

$$\pi(\theta, \lambda_{1:T}, h_{1:T}|y_{1:T}) \;\propto\; f(y_{1:T}, \lambda_{1:T}, h_{1:T}|\theta) \cdot \pi(\theta) \cdot \mathbb{1}\{\theta \in \Theta_D\}$$

where $f(y_{1:T}, \lambda_{1:T}, h_{1:T}|\theta)$ is the likelihood function, and $\mathbb{1}\{\theta \in \Theta_D\}$ is an indicator function that equals one if θ is in the determinacy region Θ_D and zero otherwise.

Conceptual simplicity aside, sampling this posterior distribution is computationally challenging. These computational challenges are magnified in larger dimensional DSGE models. For this reason, in our view, there is an urgent need for a simple toolbox that makes the fitting of such models possible, without any attendant set-up costs. The DGSE-SVt MATLAB toolkit is written to fulfill this need.

[3] It is straightforward to introduce, as in Cúrdia et al. (2014), an independent Student-t distribution with different degrees of freedom for each shock innovation. For exhibition ease, we do not consider this generalization.

3 Prior Construction

In the Bayesian estimation of DSGE models, an informative prior distribution (such as those on the policy parameters ϕ_π, γ_g, γ_z—see Appendix "A") can play an important role in emphasizing the regions of the parameter space that are economically meaningful. It can also introduce curvature into the posterior surface that facilitates numerical optimization and MCMC simulations.

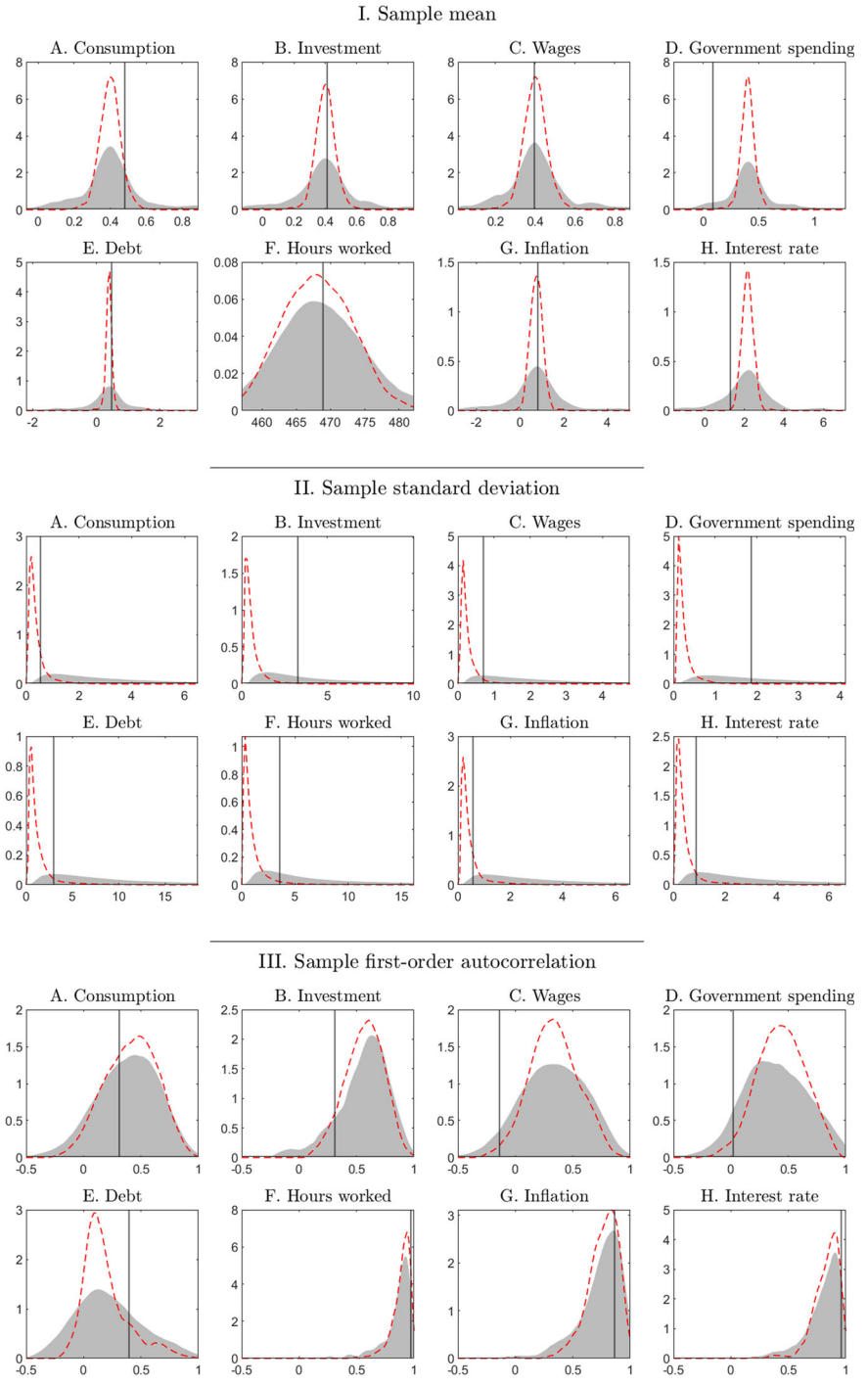
When it comes to high dimensions, however, constructing an appropriate prior becomes increasingly difficult due to the complex mapping from the structural parameters to those of the state space form. Consequently, a reasonable prior for the structural parameters may still imply a distribution of the data that is strongly at odds with actual observations. For instance, Fig. 1 shows the implied distributions for selected sample moments under the original regime-M prior and model specification of Leeper et al. (2017) (red dashed lines). Most notably, this prior places little or no mass in the neighborhood of the actual mean of government spending and the actual standard deviations of investment, government spending, debt, and hours worked (vertical lines). After taking the model to data, we also find that the posterior mass for several parameters (e.g., the habit parameter h, the nominal rigidity parameters ω_p and ω_w, and the government spending persistence ρ_g) lies entirely in the far tail of the corresponding prior, thereby introducing fragility to the inferences. To cope with these issues, we suggest a two-step approach for constructing the prior that can avoid such prior-sample conflict.

3.1 Sampling the Prior

The first step follows the sampling the prior approach in, e.g., Geweke (2005) and Chib and Ergashev (2009). In particular, one starts with a standard prior for the structural parameters. Here we take that to be the prior in Leeper et al. (2017), reproduced in Table 5 of Appendix "A".[4] Alongside, one specifies an initial prior for the volatility parameters θ^V, say one that implies a fairly persistent volatility process for each shock innovation. Then, one samples this joint prior a large number of times (say 10,000). For each parameter draw $\theta^{(g)}$, $g = 1, \ldots, G$, from the prior, under which the model has a unique bounded solution, one simulates a sample of T observations $y_{1:T}^{(g)}$. Finally, one computes the implied distributions of various functions of the data (such as the sample mean, standard deviation, and autocorrelation) and one checks whether these are close to the corresponding values in the actual data. If not, one adjusts some or all marginal components of the prior and repeats the above process.[5]

[4] Because some parameters are held fixed under each regime, effectively, θ has 49 elements and θ^S has 25 elements.

[5] In the Leeper et al. (2017) setting with Gaussian shocks and constant volatility, this step suggests that the original prior for the standard deviation parameters should be adjusted. Alternatively, one could also adjust other components of the prior for θ^S.


I. Sample mean
A. Consumption
B. Investment
C. Wages
D. Government spending
E. Debt
F. Hours worked
G. Inflation
H. Interest rate
II. Sample standard deviation
A. Consumption
B. Investment
C. Wages
D. Government spending
E. Debt
F. Hours worked
G. Inflation
H. Interest rate
III. Sample first-order autocorrelation
A. Consumption
B. Investment
C. Wages
D. Government spending
E. Debt
F. Hours worked
G. Inflation
H. Interest rate

◀**Fig. 1** Distributions of simulated selected quantities obtained by sampling the prior, and then the outcomes given drawings from the prior. Notes: This figure is based on regime-M model. Each panel compares the resulting densities under Gaussian shocks with constant volatility (red dashed line) with that under Student-t shocks with stochastic volatility (shaded area). The kernel smoothed densities are estimated using 10, 000 simulated draws. Vertical lines denote the real data counterparts

It is clear from Fig. 1 that under the adjusted prior, reported in Table 5 of Appendix "A", the Leeper et al. (2017) model extended with Student-t shocks and stochastic volatility implies distributions of the data that capture the corresponding real data quantities in their relatively high density regions (shaded areas).

3.2 Training Sample Prior

In the second step, given the adjusted prior from the first step, one uses the TaRB-MH algorithm to estimate the DSGE model on the initial 50 observations running from 1955:Q1 to 1967:Q2. The posterior draws from this run are used to form the prior. Specifically, the prior type of each parameter is left unchanged, but its location (dispersion) are set to the corresponding mean (twice standard deviation). At this point, we suggest that each location parameter μ of the volatility process be assigned a Student-t distribution with 2.1 degrees of freedom. This two-step construction tends to avoid stark conflict between the prior and the likelihood.

4 Posterior Sampling

We use two primary steps to sample the posterior distribution. The first step samples the 25 structural parameters in θ^S from the conditional posterior $\pi(\theta^S|y_{1:T}, \theta^V, \lambda_{1:T}, h_{1:T})$ by the TaRB-MH algorithhm of Chib and Ramamurthy (2010). The second step samples the remaining blocks, including the 24 volatility parameters in θ^V, the 166 non-Gaussian latent variables in $\lambda_{1:T}$, and the 1,328 nonlinear latent variables in $h_{1:T}$, from the conditional posterior $\pi(\theta^V, \lambda_{1:T}, h_{1:T}|y_{1:T}, \theta^S)$ by the Kim et al. (1998) method. Iterating the above cycle until convergence produces a sample from the joint posterior $\pi(\theta, \lambda_{1:T}, h_{1:T}|y_{1:T})$. We provide a brief summary of these steps and refer readers to the original papers for further details.

4.1 Sampling Structural Parameters

The first step entails sampling θ^S from

$$\pi(\theta^S|y_{1:T}, \theta^V, \lambda_{1:T}, h_{1:T}) \propto f(y_{1:T}|\theta^S, \lambda_{1:T}, h_{1:T}) \cdot \pi(\theta^S) \cdot \mathbb{1}\{\theta \in \Theta_D\}$$

using the TaRB-MH algorithm. To form a random partition $\theta^S = (\theta_1^S, \ldots, \theta_B^S)$, we initialize θ_1^S with the first element from a permuted sequence of θ^S, and start a new

block with every next element with probability $1-p$. As a result, the average size of a block is given by $(1-p)^{-1}$. In our benchmark setting, we set $p = 0.7$ so that each block contains three to four parameters on average. To generate a candidate draw, we tailor the Student-t proposal density to the location and curvature of the posterior distribution for a given block using the BFGS quasi-Newton method (available as a MATLAB function `csminwel` written by Chris Sims).[6]

We also introduce a new procedure, i.e., tailoring-at-random-frequency, to accelerate the TaRB-MH algorithm. The idea is similar in essence to grouping the structural parameters into random blocks. Because the tailored proposal density in the current iteration may remain efficient for the next few iterations, there is typically no need to re-tailor the proposal density in every iteration. Nevertheless, there is still a chance that the re-tailored proposal density will be quite different from the recycled one. Therefore, randomizing the number of iterations before new blocking and tailoring ensures that the proposal density remains well-tuned on average. The reciprocal of this average number, which we call the tailoring frequency ω, as well as a number of optional user inputs (e.g., the blocking probability p), can be specified flexibly in the program `tarb.m`. In our benchmark setting, we set $\omega = 0.5$ so that each proposal density is tailored every second iteration on average.

4.2 Sampling Latent Variables and Volatility Parameters

The second step involves augmenting the remaining blocks with 1,328 shock innovations $\epsilon_{1:T}$ and then sampling the joint posterior $\pi(\theta^V, \epsilon_{1:T}, \lambda_{1:T}, h_{1:T}|y_{1:T}, \theta^S)$. To this end, Gibbs sampling is applied to the following conditional densities

$$\pi(\epsilon_{1:T}|y_{1:T}, \theta, \lambda_{1:T}, h_{1:T}), \qquad \pi(\lambda_{1:T}|y_{1:T}, \theta, \epsilon_{1:T}, h_{1:T}), \qquad \pi(\theta^V, h_{1:T}|y_{1:T}, \theta^S, \epsilon_{1:T}, \lambda_{1:T})$$

The first density is sampled with the disturbance smoother of Durbin and Koopman (2002). The second density is sampled as in Chib and Ramamurthy (2014) based on a mixture normal representation of the Student-t distribution. We invest in the most efficient way of sampling the last density by the integration sampler of Kim et al. (1998).

4.3 Results

We apply the above steps as coded in our MATLAB toolkit to estimate the high-dimensional DSGE model based on the post-training sample of 166 quarterly observations from 1967:Q3 to 2008:Q4. With the ultimate goal of forecasting in mind, we present the estimation results for the model of best fit among all

[6] The same optimization procedure is applied to obtain a starting value $\theta^{S,(0)}$ for the chain. This procedure is repeated multiple times, each of which is initialized at a high density point out of a large number of prior parameter draws. The optimization results will be stored in the MATLAB data file `chaininit.mat`, which is saved to the subfolder `user/ltw17`.

competing specifications. This specification stands out from an extensive model search based on a marginal likelihood comparison, as described in the next section. It has regime-M in place and features heavy-tailed shocks with 5 degrees of freedom and persistent volatilities.

Because the TaRB-MH algorithm is simulation efficient, a large MCMC sample is typically not required. We consider a simulation sample size of 11,000 draws, of which the first 1,000 draws are discarded as the burn-in phase. Figure 2 provides a graphical comparison of the prior and posterior distributions of each structural parameter. The Bayesian learning is clear from the graphs. In particular, the data imply quite high habit formation and relatively high degrees of price and wage stickiness. See also Table 6 of Appendix "A" for a detailed summary of the posterior parameter estimates.

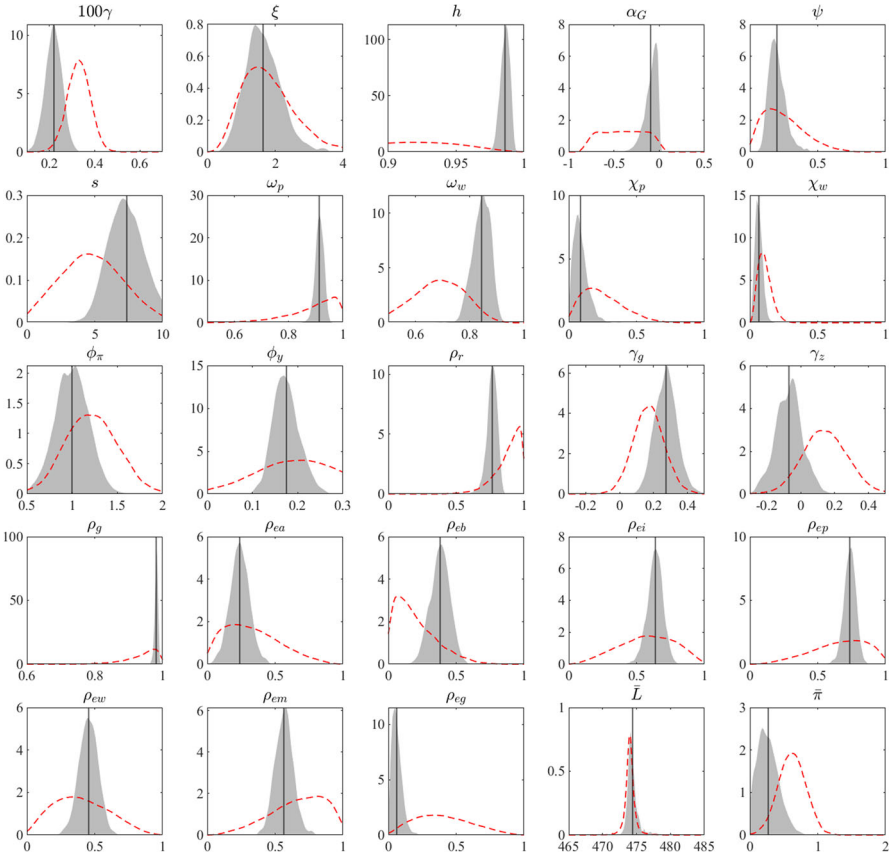


Fig. 2 Marginal prior and posterior distributions of each structural parameter. Notes: This figure is based on regime-M model. Each panel compares the prior (red dashed line) with the posterior (shaded area). Vertical lines denote posterior means. The kernel smoothed posterior densities are estimated using 10, 000 TaRB-MH draws

Figure 3 plots the estimated historical log-volatility series for 1967:Q3 to 2008:Q4. Overall, these estimates display clear countercyclical time variation, with pronounced increases in volatility accompanying the recessions. For several shock innovations, volatility becomes lower by historical standards since the 1980s so that the Great Moderation is also evident.

To see the sampling efficiency of the TaRB-MH algorithm, it is informative to examine the serial correlation among the sampled draws. Figures 8 and 9 of Appendix "A" display the autocorrelation function for each element of θ. As can be observed, the serial correlations for most parameters decay quickly to zero after a few lags. Another useful measure of the sampling efficiency is the so-called inefficiency factor, which approximates the ratio between the numerical variance of the estimate from the MCMC draws and that from the hypothetical *i.i.d.* draws. An efficient sampler produces reasonably low serial correlations and hence inefficiency

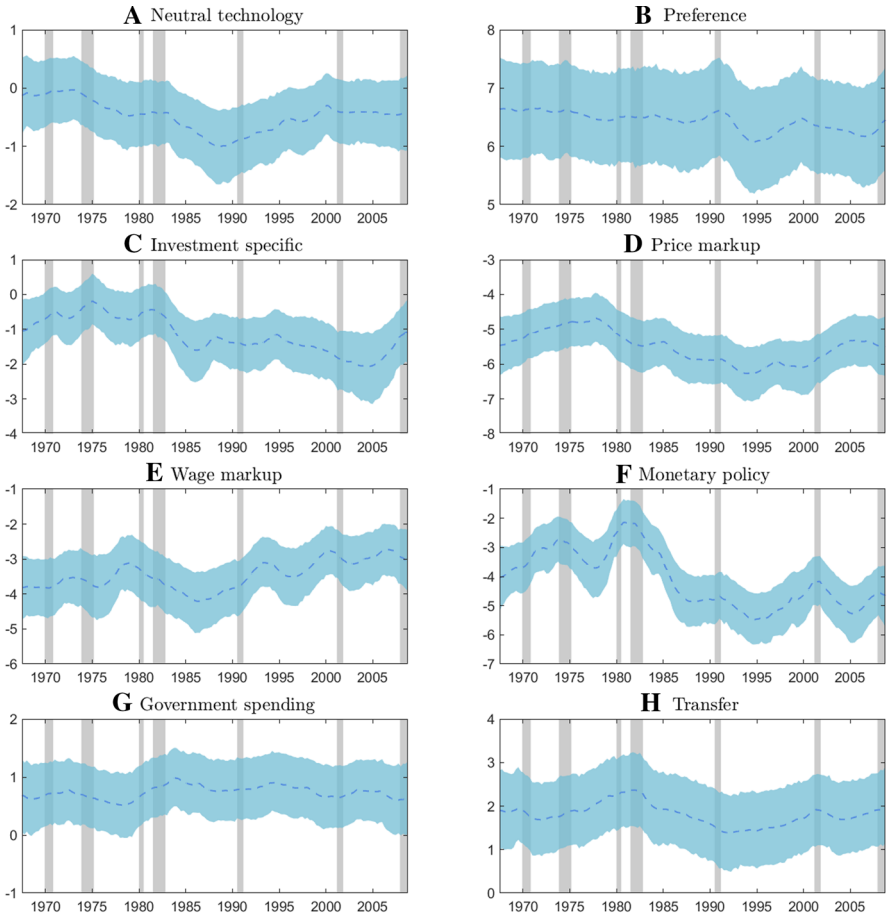


Fig. 3 Stochastic volatility of each shock innovation. Notes: Blue dashed lines denote median estimates, while blue shaded areas delineate 90% highest posterior density bands. Vertical bars indicate recessions as designated by the National Bureau of Economic Research

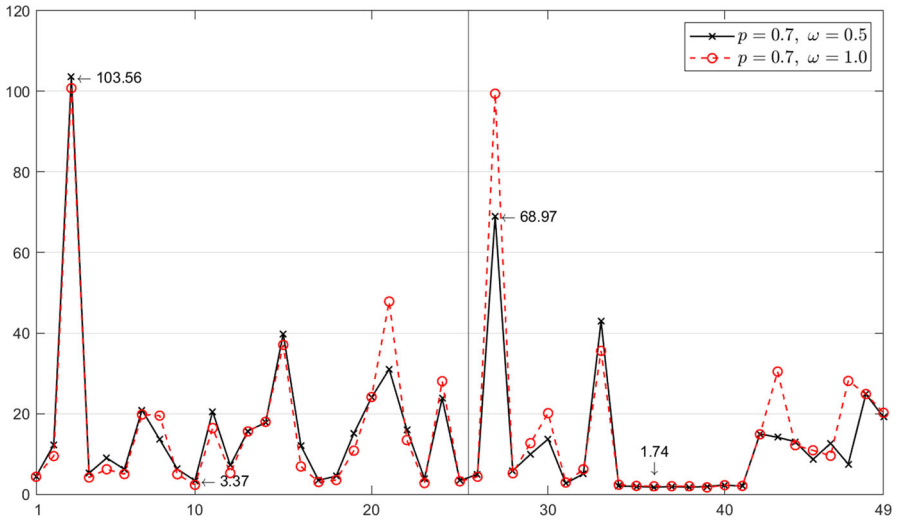


Fig. 4 Inefficiency factor of each parameter. Notes: The horizontal axis indicates parameter indices. The vertical line separates the structural (indexed by 1–25) and volatility (indexed by 26–49) parameters

factors. Figure 4 compares the inefficiency factors resulting from two different sampling schemes, each corresponding to a choice of the tailoring frequency $\omega \in \{0.5, 1.0\}$ with the same blocking probability $p = 0.7$. Compared to the more efficient setting that tailors in every iteration ($\omega = 1.0$), our benchmark setting that tailors on average every second iteration ($\omega = 0.5$) leads to very similar inefficiency factors, ranging from 3.37 (1.74) to 103.56 (68.97) with most values below 20 (15) for the structural (volatility) parameters. In conjunction with a rejection rate of approximately 50% in the M-H step, the small inefficiency factors suggest that the chain mixes well. Moreover, choices of ω have no material effect on the sampling efficiency for most volatility parameters due to the use of the integration sampler that is tailored in every iteration.[7]

While single blocking or infrequent tailoring can greatly reduce the overall runtime, it may also add considerably to the average inefficiency factor. Therefore, in practice, we suggest setting $p \in [0.6, 0.9]$ and $\omega \in [0.2, 1.0]$ to maintain a good balance between runtime and simulation efficiency.

4.4 Simulation Evidence

We also estimate the same high-dimensional DSGE model based on a simulated data set that is generated under fat-tailed shocks with 15 degrees of freedom and

[7] It is worth noting that the total runtime for generating 11,000 draws is about 30 hours for $\omega = 0.5$ and is about 43 hours for $\omega = 1.0$. These numbers were measured based on a computer with Intel (R) i5-6500 CPU with 16 GB RAM. With an average inefficiency factor of 14.46 for $\omega = 0.5$ and 15.66 for $\omega = 1.0$, these numbers translate into 25.36 and 16.34 *i.i.d.*-equivalent draws per hour, respectively. This comparison exemplifies that the tailoring-at-random-frequency feature can effectively improve the computational efficiency of the TaRB-MH algorithm.

persistent volatilities. We set the sample size to 200, which is meant to be 50 years of quarterly observations, and use the initial 50 observations to construct a training sample prior. Table 5 of Appendix "A" lists the parameter values used for the data generating process under regime-M (column 'DGP').[8] Figure 10 provides a graphical comparison of priors and posteriors. For most parameters, the posterior mass concentrates around the corresponding true value. Figure 11 further reveals that the estimated log-volatility series largely captures the level and all major trends of the true series for each shock innovation. These plots are relegated to Appendix "A".

5 Marginal Likelihood

Given the output of the efficient TaRB-MH algorithm, we suggest calculating the marginal likelihood by the method of Chib (1995), as modified for M-H chains in Chib and Jeliazkov (2001). This method is computed via the identity

$$m(y_{1:T}|\mathscr{M}) = \frac{\frac{1}{c} f(y_{1:T}|\mathscr{M},\theta)\cdot\pi(\theta|\mathscr{M})\cdot\mathbb{1}\{\theta\in\Theta_D\}}{\pi(\theta|\mathscr{M},y_{1:T})}$$

where $\mathscr{M}$ denotes the model label, $c = \int_{\theta\in\Theta_D}\pi(\theta|\mathscr{M})d\theta$, and the right-hand-side terms are evaluated at a single high density point θ^*. We obtain the likelihood ordinate by a mixture version of the Kalman filter introduced by Chen and Liu (2000), as facilitated by the conditionally Gaussian and linear structure of the DSGE model solution. In our application, we find that 10,000 particles are sufficient to deliver a robust estimate of $f(y_{1:T}|\mathscr{M},\theta^*)$. We obtain the high-dimensional ordinate in the denominator after decomposing it as

$$\begin{aligned}\pi(\theta^*|\mathscr{M},y_{1:T}) = \pi(\theta_1^*|\mathscr{M},y_{1:T})\cdot\pi(\theta_2^*|\mathscr{M},y_{1:T},\theta_1^*)\cdots\\ \pi(\theta_B^*|\mathscr{M},y_{1:T},\theta_1^*,\ldots,\theta_{B-1}^*)\end{aligned}$$

where B refers to the number of blocks (that is under our control), and then estimate each of these reduced ordinates from the MCMC output of reduced runs.

An interesting point is that these reduced runs are independent of each other and can be done in parallel. Thus, all reduced ordinates can be estimated at the cost of one reduced run, regardless of the size of B. This parallel computation is built into our MATLAB toolkit. In our application, we set the total number of blocks to $B = 15$, including seven almost equally sized blocks for θ^S arranged first, followed by eight blocks (μ,ϕ,ω^2) for θ^V. All ordinates are then simultaneously estimated using MATLAB's multi-core processing capacity via its Parallel Computing Toolbox.

[8] The DGPs for the structural parameters correspond to their full sample (1955:Q1–2014:Q2) regime-M estimates reported in Leeper et al. (2017).

5.1 Reliability

We recommend the Chib and Jeliazkov (2001) method because it is reliable and because other methods do not generalize to our high-dimensional DSGE models with non-Gaussian and/or nonlinear latent variables.[9] As shown in Chib and Ramamurthy (2010), efficient MCMC estimation automatically delivers an efficient estimate of the conditional posterior ordinate $\pi(\theta_b, \ldots, \theta_B, \lambda_{1:T}, h_{1:T}|\mathscr{M}, y_{1:T}, \theta_1^*, \ldots, \theta_{b-1}^*)$ from the output of the reduced MCMC simulation in which θ_b is a fixed block and the remaining structural parameters, if any, form random blocks.[10] This, in turn, is used to estimate the marginal likelihood in logarithm

$$\log m(y_{1:T}|\mathscr{M}) = -\log c + \log f(y_{1:T}|\mathscr{M}, \theta^*) + \log \pi(\theta^*|\mathscr{M}) - \sum_{b=1}^{B} \log \pi(\theta_b^*|\mathscr{M}, y_{1:T}, \theta_1^*, \ldots, \theta_{b-1}^*)$$

Figure 5 displays the sequence of posterior ordinate and marginal likelihood estimates from the best fit model, as functions of the number of MCMC draws, for efficient and (relatively) less efficient TaRB-MH implementations. These estimates settle down quickly (after say 1,000 draws are made) and converge to the same limit point, leading to an estimated log marginal likelihood of about -1579.65 with a numerical standard error of about 0.12. This underscores the point that, since the Chib (1995) method is underpinned by whatever MCMC algorithm is used in the posterior simulation, the efficiency of the MCMC simulator is germane to the calculation of the marginal likelihood.

5.2 Regime Comparison

Because regimes M and F of the Leeper et al. (2017) model imply completely different mechanisms for price level determination and therefore different policy advice, identifying which policy regime produced the real data is key to making good policy choices. While it is difficult to explore the entire model space, we perform extensive regime comparisons by estimating the marginal likelihood for both regimes with four choices of the degrees of freedom $\nu \in \{2.1, 5, 15, 30\}$ and three choices of the volatility persistence $\phi \in \{0.1, 0.5, 0.95\}$. The resulting model space contains a total of 24 relevant models that are simultaneously confronted with

[9] For instance, the modified harmonic mean (MHM) estimator of Gelfand and Dey (1994), used, for example, in Justiniano and Primiceri (2008) and Cúrdia et al. (2014) in medium-scale DSGE models with Student-t shocks and stochastic volatility, always favors a model specification with stronger latent features, e.g., shocks with fatter tails or volatilities with more persistence. This extreme result emerges even when the true model exhibits weak evidence of these features, such as those considered in Sect. 5.3.

[10] In contrast, Justiniano and Primiceri (2008, p. 636) and Herbst and Schorfheide (2016, p. 97) estimate the posterior ordinate in a single block, with the random-walk M-H, both detrimental to getting reliable and efficient marginal likelihood estimates, as already documented in Chib and Jeliazkov (2001).

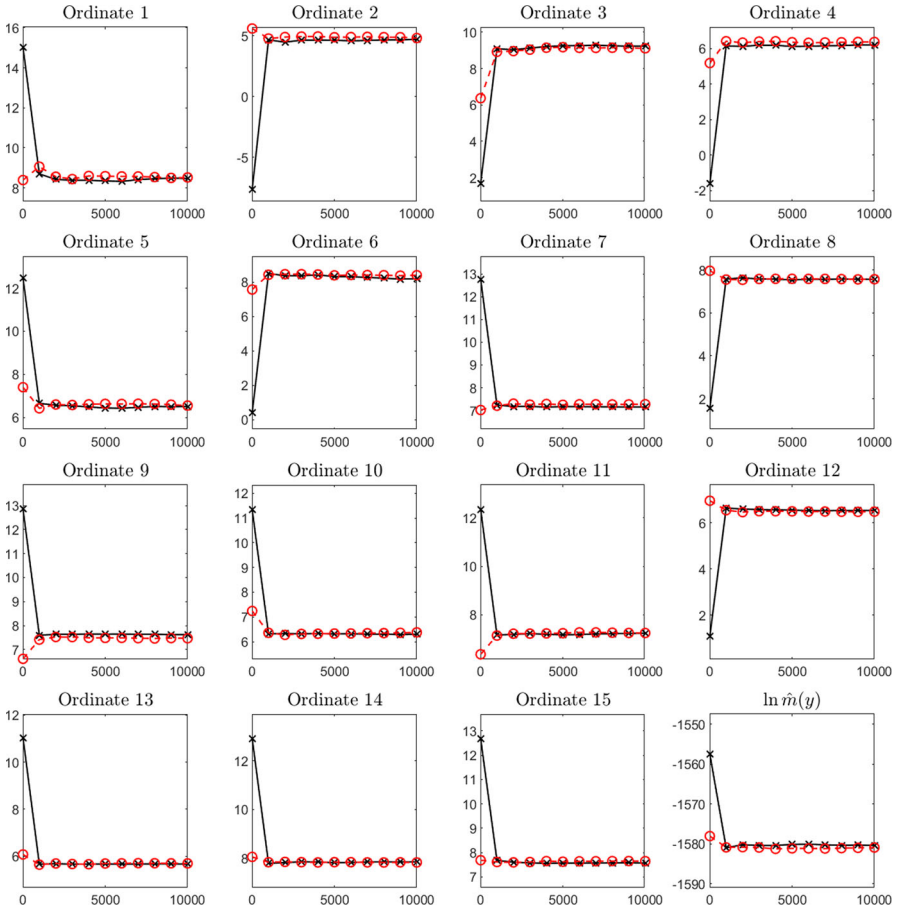


Fig. 5 Recursive posterior ordinates and marginal likelihood. Notes: Ordinates 1–7 (8–15) correspond to structural (volatility) parameters. The last panel depicts the estimated marginal likelihood. Black solid lines (with cross marker) correspond to the benchmark setting ($p = 0.7$, $\omega = 0.5$). All estimates are in logarithm scale

the data over the period from 1967:Q3 to 2008:Q4, similar in spirit to the Bayesian model scan framework proposed by Chib and Zeng (2020).[11]

Two aspects of the marginal likelihood estimates reported in Table 1 are worth highlighting. First, the data systematically prefer regime-M over regime-F in all cases, which corroborates the regime ranking found by Leeper et al. (2017) with Gaussian shocks and constant volatility. The small numerical standard errors point to the numerical accuracy of the marginal likelihood estimates. Second, reading the table by row (column) for each regime suggests that the data exhibit quite strong evidence in favor of heavy-tailed shocks (persistent volatility process). Indeed, each

[11] All computations performed in this section are executed on the High Performance Computing Cluster maintained by Saint Louis University (https://slu.zendesk.com/hc).

Table 1 Log marginal likelihood estimates

ν	$\phi = 0.1$ (weak)		$\phi = 0.5$ (moderate)		$\phi = 0.95$ (strong)	
	M	F	M	F	M	F
30 (light)	−1640.73	−1650.03	−1627.24	−1638.81	−1597.72	−1609.69
	(0.15)	(0.15)	(0.14)	(0.15)	(0.12)	(0.13)
15 (fat)	−1622.26	−1631.66	−1612.62	−1624.22	−1586.70	−1596.68
	(0.14)	(0.14)	(0.13)	(0.13)	(0.12)	(0.13)
5 (heavy)	−1605.77	−1616.95	−1600.18	−1611.05	−1579.65	−1593.11
	(0.15)	(0.14)	(0.14)	(0.13)	(0.12)	(0.12)
2.1 (heavy)	−1622.31	−1629.38	−1618.37	−1630.84	−1602.76	−1615.60
	(0.15)	(0.15)	(0.14)	(0.12)	(0.11)	(0.12)

Numerical standard errors are reported in parentheses. All estimates are obtained using 15 reduced TaRB-MH runs under the benchmark setting ($p = 0.7$, $\omega = 0.5$), including 7 runs for the structural parameters and 8 runs for the volatility parameters. 10,000 posterior draws are made for each reduced run

feature is important for improving the fit, even after accounting for the other, and the model that fits best is regime-M with $\nu = 5$ and $\phi = 0.95$.

5.3 Simulation Evidence

This section furnishes additional evidence that demonstrates the reliability of the Chib (1995) method. For each regime, we generate 20 data sets of 100 quarterly observations using the subsample parameter estimates reported in Leeper et al. (2017), which are also reproduced in Table 7 of Appendix "A". We then estimate three versions of each regime model that differ in the volatility specification. Based on the marginal likelihood estimates, we count the number of times that each of the six regime-volatility specifications is picked across the 20 simulated data sets. Table 2 summarizes the simulation results.

The first data generating process assumes that regime-M is in place and the shock innovations follow a multivariate Student-t distribution with fat tails, i.e., $\nu = 15$,

Table 2 Number of picks for each model specification

DGP 1: regime-M with $\nu = 15$			DGP 2: regime-F with $\phi = 0.5$		
ν	regime-M	regime-F	ϕ	regime-M	regime-F
30 (light)	4	0	0.1 (weak)	0	9
15 (fat)	15	0	0.5 (moderate)	0	10
5 (heavy)	1	0	0.9 (strong)	0	1

The shock innovations have constant volatility under DGP 1 and follow Gaussian distribution under DGP 2. The number of simulations performed for each DGP is 20

and constant volatility. For each regime, we fit the model with three degrees of freedom: $\nu = 30$ (light), $\nu = 15$ (fat), and $\nu = 5$ (heavy). As can be seen from the left panel of Table 2, the correct degrees of freedom is picked 15 times and the correct policy regime is always picked. We have also computed the marginal likelihood by the MHM method as implemented in Justiniano and Primiceri (2008) and Cúrdia et al. (2014). We find that nearly all data sets favor the lowest degrees of freedom, i.e., $\nu = 5$.

The second data generating process assumes that regime-F is in place and the shock innovations follow a multivariate Gaussian distribution with moderate time-varying volatility, i.e., $\phi = 0.5$. For each regime, we fit the model with three degrees of persistence in volatility: $\phi = 0.1$ (weak), $\phi = 0.5$ (moderate), and $\phi = 0.9$ (strong). As shown in the right panel of Table 2, with only one exception, the data overwhelmingly favor weak to moderate degree of persistence in volatility under the true regime, which is preferred by all data sets over the alternative regime. However, the computation based on the MHM method always overestimates the importance of stochastic volatility and selects $\phi = 0.9$. This result emerges despite the fact that all data sets are relatively short-lived and generated by a model with 'close' to constant volatility process.

6 Prediction

Because a good understanding of the current and future state of the economy is essential to develop and implement sound economic policies, generating a predictive distribution for the future path of the economy constitutes an important part of the policy analysis. To facilitate this goal, our MATLAB toolkit also produces, as a byproduct of the efficient TaRB-MH algorithm and the marginal likelihood computation by the Chib (1995) method, the joint predictive distribution for all observable variables at any forecasting horizon. For illustration purposes, Sect. 6.1 presents such a predictive distribution based on the best fitting model that is selected by the marginal likelihood comparison. Using the predictive distribution for wages as an example, Sect. 6.2 highlights the importance of allowing for non-Gaussian structural shocks with time-varying variances in the context of out-of-sample prediction. Finally, Sect. 6.3 evaluates the predictive performance by comparing the accuracy of point and density forecasts between a small-scale DSGE model and our high-dimensional DSGE model.

6.1 Sampling the Predictive Distribution

Let $y_{1:T}$ be the data used to perform estimation, inference, and model selection. In addition, denote $y_{T+1:T+h}$ the future path of the observable variables in the model economy. Then the predictive distribution is defined as

$$p(y_{T+1:T+h}|y_{1:T}) = \int p(y_{T+1:T+h}|y_{1:T}, \theta) \cdot p(\theta|y_{1:T}) d\theta$$

where the above integration is numerically approximated by first sampling the posterior $p(\theta|y_{1:T})$ a large number of times by the TaRB-MH algorithm and then simulating a future path $y^{(g)}_{T+1:T+h}$ for each parameter draw. This amounts to moving model variables forward with θ and $y_{1:T}$. We call $p(y_{i,T+h}|y_{1:T})$ the h-step-ahead predictive distribution for the ith variable generated in period T.[12]

Now we generate the one-quarter-ahead predictive distribution for all eight observable variables based on the best fitting model as measured by the marginal likelihood. Throughout the entire forecasting horizon, this model operates under regime-M and has Student-t shocks with stochastic volatilities. The first predictive distribution is generated using observations from the third quarter of 1967 to the fourth quarter of 2008, which is about six months before the Business Cycle Dating Committee of the National Bureau of Economic Research announces the end of the Great Recession. The forecasting horizon starts from the first quarter of 2009 and ends at the second quarter of 2014, covering the whole economic recovery period from the Great Recession.

Figure 6 displays the median forecasts with 90% credible bands computed from the predictive distribution of regime-M over the full forecasting horizon. Overall the model performs quite well in tracking the recovery path of most observable variables.

6.2 Importance of Non-Gaussian Shocks

As the marginal likelihood comparison reveals, one needs a flexible way to model structural shocks in the model economy to explain the U.S. macroeconomic variables. The need of flexible distributional assumptions, such as Student-t shocks with stochastic volatility, can also be seen from our generated predictive densities as well. The left panel of Fig. 7 plots the 90% credible sets for wages based on two predictive distributions: one under Gaussian shocks with constant variance and another under Student-t shocks with time-varying variance. It is noticeable that the uncertainty bands are much wider for the model under Student-t shocks with time-varying variance. To understand this stark difference, the right panel of Fig. 7 plots the time series of wages over the full sample. As pointed out by Champagne and Kurmann (2013), wages in the U.S. have become more volatile over the past 20 years. For example, the standard deviation of wages was 0.55 between 1955:Q1 and 1999:Q4, and 1.05 between 2000:Q1 and 2014:Q2. The heightened volatility of wages after 2000 is captured by the model with stochastic volatility, which adaptively widens the predictive distribution for wages. On the other hand, the model with constant variance misses this important change in volatility. In turn, its predictive distribution of wages is too narrow, underestimating the uncertainty in the future path of wages. In general, allowing for time-varying volatility produces similar improvements in the quality of DSGE-based interval and density forecasts [see, e.g., Diebold et al. (2017)]. Thus, we expect that our toolbox, by making it easy to incorporate non-Gaussian errors and time-varying variances, will be useful

[12] This density function, when it is evaluated at the realized value of $y_{T+1:T+h}|y_{1:T}$, is called a predictive likelihood. It is oftentimes used for comparing different models [see, e.g., Chib and Greenberg (1995a)].

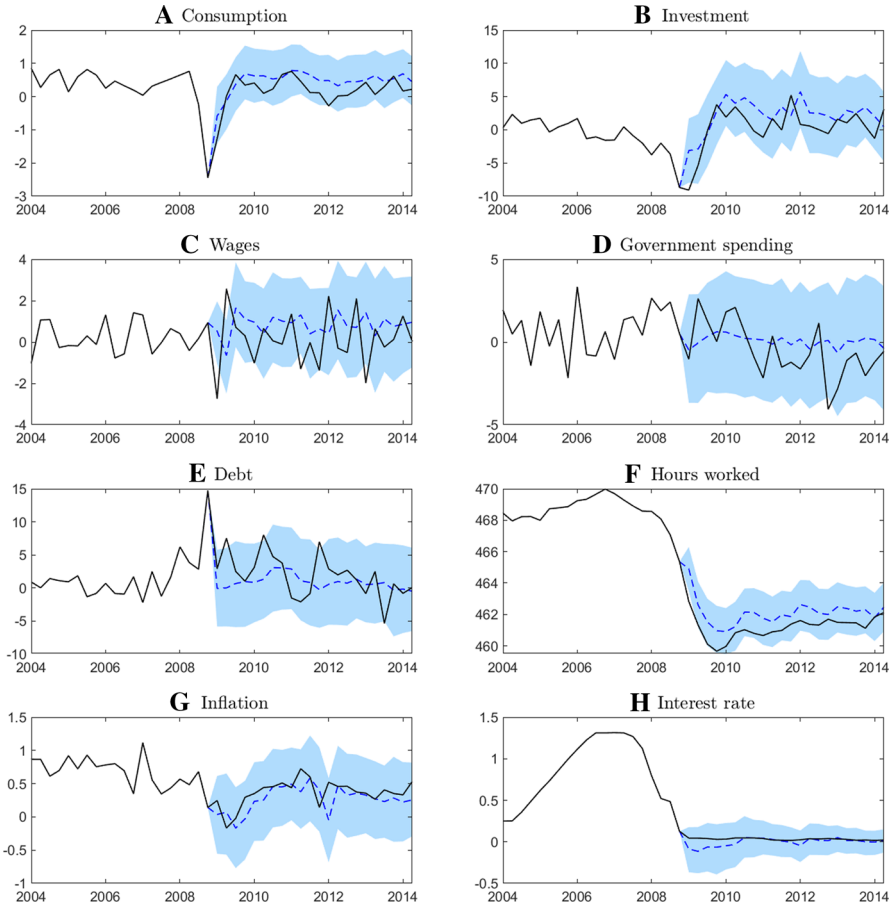


Fig. 6 DSGE-model forecast of each observable variable. Notes: Each panel compares the one-quarter-ahead posterior forecast of regime M with real data (black solid lines). Blue dashed lines denote median forecasts, while blue shaded areas delineate 90% highest predictive density bands

for researchers and policymakers interested in better out-of-sample performance of DSGE models.

6.3 Predictive Performance Comparison

Although regime-M yields a higher marginal likelihood relative to regime-F, one may still be interested in knowing how the two policy regimes compare in terms of the quality of point and density forecasts over the forecasting horizon. It is also interesting to compare the forecasts from a medium-scale DSGE model with those from a small-scale one when both models are equipped with Student-t shocks and stochastic volatility. Specifically, we compare the point and density forecasts generated from regimes M and F, and a small-scale DSGE model described in Appendix "B". Starting from the first quarter of 2009, we recursively estimate the

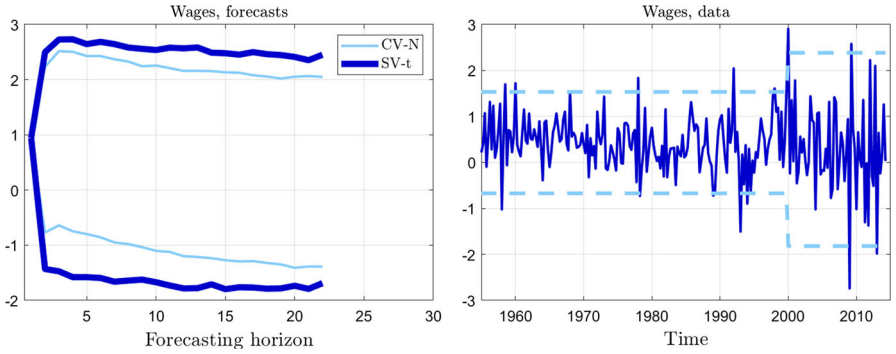


Fig. 7 Predictive distribution and data for wages. Notes: Predictive distributions are constructed using data up to 2008:Q4. The one-step-ahead prediction corresponds to 2009:Q1. The left panel plots 90% prediction intervals of regime-M under Gaussian shocks with constant variance (labeled 'CV-N', thick line) and Student-t shocks with time-varying variance (labeled 'SV-t', thin line). The right panel plots the time series of wages (solid line). Dashed lines delineate two standard deviations from the mean for two sub-samples, i.e., pre- and post-2000

three models and generate one-quarter-ahead to two-year-ahead point and density forecasts until the second quarter of 2014, which results in 22 quarters of evaluation points for the one-quarter-ahead prediction. Since the small-scale model contains fewer observable variables, our evaluation exercise only considers the common set of observable variables: consumption growth, inflation rate, and federal funds rate. The aim of this comparison is to get information about the strengths and weaknesses of DSGE model elaborations.

In each model, for each observable variable and forecasting horizon, the point prediction is the mean of the corresponding predictive distribution. Let $\widehat{y}_{i,t+h|t}$ denote the h-step-ahead point prediction for the ith variable generated at time t. To compare the quality of point forecasts, we report the root mean squared error (RMSE) for the point prediction

$$\mathrm{RMSE}(\widehat{y}_{i,t+h|t}, y_{i,t+h}) = \sqrt{\frac{1}{22-h}\sum_{t=2009:\mathrm{Q1}}^{2014:\mathrm{Q2}-\mathrm{h}}\left(y_{i,t+h}-\widehat{y}_{i,t+h|t}\right)^2}$$

where 2014:Q2$-h$ denotes h-quarters before 2014:Q2 and $y_{i,t+h}$ is the actual value for the ith variable at time $t+h$. The model with a smaller RMSE is preferred as the smaller forecast error is desirable. To compare the precision of predictive densities, we compute the continuous ranked probability score (CRPS), which is defined as

$$\mathrm{CRPS}(F_{i,t+h|t}(z), y_{i,t+h}) = \int_{\mathbb{R}}\left(F_{i,t+h|t}(z) - 1\{y_{i,t+h} \le z\}\right)^2 dz$$

where $F_{i,t+h|t}(z)$ is the h-step-ahead predictive cumulative distribution of the ith variable generated at time t. The CRPS is one of the proper scoring rules, and the predictive distribution with a smaller CRPS is preferred as this measure can be viewed as the divergence between the given predictive distribution and the

Table 3 Point forecast comparison, RMSE

Model	$h = 1$Q	$h = 2$Q	$h = 4$Q	$h = 8$Q
(a) Consumption growth				
Small-scale	0.32	0.28	0.25	0.28
Regime-M	0.44 (0.06)	0.48 (0.20)	0.50 (0.22)	0.48 (0.12)
Regime-F	0.40 (0.23)	0.39 (0.16)	0.36 (0.38)	0.37 (0.11)
(b) Inflation rate				
Small-scale	0.26	0.32	0.46	0.58
Regime-M	0.24 (0.40)	0.28 (0.31)	0.37 (0.12)	0.44 (0.04)
Regime-F	0.34 (0.00)	0.53 (0.08)	0.86 (0.10)	1.14 (0.14)
(c) Federal funds rate				
Small-scale	0.21	0.38	0.64	0.94
Regime-M	0.06 (0.00)	0.12 (0.01)	0.19 (0.01)	0.42 (0.01)
Regime-F	0.06 (0.00)	0.12 (0.02)	0.18 (0.02)	0.22 (0.01)

Each entry reports the RMSE based on the point forecast with the *p*-value of Diebold-Mariano (DM) tests of equal MSE in parentheses, obtained using the fixed-*b* critical values. The standard errors entering the DM statistics are computed using the equal-weighted cosine transform (EWC) estimator with the truncation rule recommended by Lazarus et al. (2018)

Table 4 Density forecast comparison, average CRPS

Model	$h = 1$Q	$h = 2$Q	$h = 4$Q	$h = 8$Q
(a) Consumption growth				
Small-scale	0.21	0.2	0.19	0.2
Regime-M	0.26 (0.08)	0.28 (0.28)	0.29 (0.31)	0.28 (0.20)
Regime-F	0.23 (0.40)	0.22 (0.41)	0.21 (0.72)	0.22 (0.46)
(b) Inflation rate				
Small-scale	0.15	0.18	0.26	0.34
Regime-M	0.14 (0.48)	0.17 (0.53)	0.23 (0.29)	0.28 (0.11)
Regime-F	0.20 (0.00)	0.31 (0.03)	0.52 (0.05)	0.69 (0.08)
(c) Federal funds rate				
Small-scale	0.13	0.24	0.43	0.67
Regime-M	0.04 (0.00)	0.07 (0.02)	0.13 (0.01)	0.27 (0.01)
Regime-F	0.04 (0.00)	0.07 (0.02)	0.12 (0.02)	0.19 (0.01)

Each entry reports the average CRPS over the evaluation period with the *p*-value of Diebold-Mariano (DM) tests of equal CRPS in parentheses, obtained using the fixed-*b* critical values. The standard errors entering the DM statistics are computed using the equal-weighted cosine transform (EWC) estimator with the truncation rule recommended by Lazarus et al. (2018)

unattainable oracle predictive distribution that puts a probability mass only on the realized value. Tables 3 and 4 report the RMSE and average CRPS, respectively, of consumption growth, inflation rate, and federal funds rate based on all three models.

Forecasts from the medium-scale models are significantly more accurate for the federal funds rate at all horizons. On the other hand, forecasts from the small-scale model are more accurate for the consumption growth at all horizons although the difference is only statistically significant at the one-quarter-ahead horizon. The major difference between regimes M and F lies in the inflation forecasts, and the model under regime-M produces forecasts with lower RMSEs (CRPSs). The RMSE (CRPS) gaps get wider as the forecasting horizon extends, and the RMSE (CRPS) from regime-M becomes more than half of that from regime-F. In contrast, the forecasts from regime-F fare slightly better for the consumption growth at all horizons and are most accurate for the federal funds rate at the two-year-ahead horizon.

In sum, there is no clear winner in this comparison. The small-scale model performs better for forecasting the consumption growth. The medium-scale model, on the other hand, performs the best under regime-M for forecasting the inflation rate but does not generate better forecasts under regime-F except for forecasting the federal funds rate in the long run. Although the evaluation period is too short-lived to draw a definite conclusion, the results from this out-of-sample forecasting exercise indicate that there is still room for improvement even for the more complex models.

7 Concluding Remarks

We have given pointers on the fitting and comparison of high-dimensional DSGE models with latent variables and shown that the TaRB-MH algorithm of Chib and Ramamurthy (2010) allows for the efficient estimation of such models. We implement the TaRB-MH algorithm for the class of DSGE models with stochastic volatility and t errors. We introduce two new features into the original TaRB-MH algorithm and demonstrate that they lead to a decent computational gain. By tailoring the proposal density at a random frequency, we speed up the posterior sampling without losing the validity and efficiency of the original algorithm. We show that the method of Chib (1995) and Chib and Jeliazkov (2001), in conjunction with a parallel implementation of the required reduced MCMC runs, can be used to obtain reliable and fast estimates of the marginal likelihood. In addition, we emphasize the importance of training sample priors, which is new in the DSGE context, and the use of the Student-t, as opposed to the normal family, as the prior distribution for location-type parameters. With the help of a user-friendly MATLAB toolkit, these methods can be readily employed in academic and central bank applications to conduct DSGE model comparisons and to generate point and density forecasts. Finally, in ongoing work, we are applying this toolkit, without modification and any erosion in performance, to open economy DSGE models that contain more than twice as many parameters and latent variables as the model showcased in this paper. Findings from this analysis will be reported elsewhere.

Appendix A: Leeper-Traum-Walker ModelLinearized System

Unless otherwise noted, we let $\hat{x_t} \equiv \ln x_t - \ln x$ denote the log-deviation of a generic variable x_t from its steady state x. We also divide a non-stationary variable X_t by the level of technology A_t and express the detrended variable as $x_t = X_t/A_t$.

Firms

The production sector consists of firms that produce intermediate and final goods. A perfectly competitive final goods producer uses intermediate goods supplied by a continuum of intermediate goods producers indexed by i on the interval [0, 1] to produce the final goods. The production technology $Y_t \leq \left(\int_0^1 Y_t(i)^{1/(1+\eta_t^p)} di\right)^{1+\eta_t^p}$ is constant-return-to-scale, where η_t^p is an exogenous price markup shock, Y_t is the aggregate demand of final goods, and $Y_t(i)$ is the intermediate goods produced by firm i.

Each intermediate goods producer follows a production technology $Y_t(i) = K_t(i)^{\alpha}\left(A_t L_t^d(i)\right)^{1-\alpha} - A_t\Omega$, where $K_t(i)$ and $L_t^d(i)$ are the capital and the amount of 'packed' labor input rented by firm i at time t, and $0<\alpha<1$ is the income share of capital. A_t is the labor-augmenting neutral technology shock and its growth rate $u_t^a \equiv \ln(A_t/A_{t-1})$ equals $\gamma > 0$ when A_t evolves along the balanced growth path. The parameter $\Omega > 0$ represents the fixed cost of production.

Intermediate goods producers maximize their profits in two stages. First, they take the input prices, i.e., nominal wage W_t and nominal rental rate of capital R_t^k, as given and rent $L_t^d(i)$ and $K_t(i)$ in perfectly competitive factor markets. Second, they choose the prices that maximize their discounted real profits. Here we introduce the Calvo-pricing mechanism for nominal price rigidities. Specifically, a fraction $0<\omega_p<1$ of firms cannot change their prices each period. All other firms can only partially index their prices by the rule $P_t(i) = P_{t-1}(i)\left(\pi_{t-1}^{\chi_p}\pi^{1-\chi_p}\right)$, where $P_{t-1}(i)$ is indexed by the geometrically weighted average of past inflation π_{t-1} and steady state inflation π. The weight $0<\chi_p<1$ controls the degree of partial indexation.

The production sector can be summarized by four log-linearized equilibrium equations in terms of six parameters $(\alpha, \Omega, \beta, \omega_p, \chi_p, \eta^p)$, seven endogenous variables $(\hat{y_t}, \hat{k_t}, \hat{L_t}, \hat{r_t^k}, \hat{w_t}, \widehat{mc}_t, \hat{\pi}_t)$, and one exogenous shock $\hat{u_t^p}$:

$$\text{Production function: } \hat{y_t} = \frac{y+\Omega}{y}\left[\alpha\hat{k_t} + (1-\alpha)\hat{L_t}\right] \tag{A.1}$$

$$\text{Capital-labor ratio: } \hat{r_t^k} - \hat{w_t} = \hat{L_t} - \hat{k_t} \tag{A.2}$$

$$\text{Marginal cost: } \widehat{mc}_t = \alpha\hat{r_t^k} + (1-\alpha)\hat{w_t} \tag{A.3}$$

$$\text{Phillips equation: } \hat{\pi}_t = \frac{\beta}{1+\beta\chi_p}\mathbb{E}_t\hat{\pi}_{t+1} + \frac{\chi_p}{1+\beta\chi_p}\hat{\pi}_{t-1} + \kappa_p\widehat{mc}_t + \hat{u}_t^p \qquad (A.4)$$

where $\kappa_p \equiv [(1-\beta\omega_p)(1-\omega_p)]/[\omega_p(1+\beta\chi_p)]$, $\hat{\eta}_t^p \equiv \ln(1+\eta_t^p) - \ln(1+\eta^p)$, $\hat{\eta}_t^p$ is normalized to $\hat{u}_t^p \equiv \kappa_p\hat{\eta}_t^p$, and $\mathbb{E}_t$ represents mathematical expectation given information available at time t.

Households

The economy is populated by a continuum of households indexed by j on the interval [0, 1]. Each optimizing household j derives utility from composite consumption $C_t^*(j)$, relative to a habit stock defined in terms of lagged aggregate composite consumption hC_{t-1}^* where $0<h<1$. The composite consumption consists of private $C_t(j)$ and public G_t consumption goods, i.e., $C_t^*(j) \equiv C_t(j) + \alpha_G G_t$, where α_G governs the degree of substitutability of the consumption goods. Each household j also supplies a continuum of differentiated labor services $L_t(j,l)$ where $l \in [0,1]$. Households maximize their expected lifetime utility $\mathbb{E}_0\sum_{t=0}^{\infty}\beta^t u_t^b\left[\ln(C_t^*(j) - hC_{t-1}^*)\right.$ $\left.-L_t(j)^{1+\xi}/(1+\xi)\right]$, where $0<\beta<1$ is the discount rate, $\xi > 0$ is the inverse of Frisch labor supply elasticity, and u_t^b is an exogenous preference shock.

Households have access to one-period nominal private bonds $B_{s,t}$ that pay one unit of currency at time $t+1$, sell at price R_t^{-1} at time t, and are in zero net supply. They also have access to a portfolio of long-term nominal government bonds B_t, which sell at the price P_t^B at time t. Maturity of these zero-coupon bonds decays at the constant rte $0<\rho<1$ to yield the average duration $(1-\rho\beta)^{-1}$. Households receive bond earnings, labor and capital rental income, lump-sum transfers from the government Z_t, and profits from firms Π_t. They spend income on consumption, investment I_t, and bonds. The nominal flow budget constraint for household j is given by

$$\begin{aligned}
&(1+\tau^C)P_tC_t(j) + P_tI_t(j) + P_t^BB_t(j) + R_t^{-1}B_{s,t}(j)\\
&\quad = (1+\rho P_t^B)B_{t-1}(j) + B_{s,t-1}(j)\\
&\quad + (1-\tau^L)\int_0^1 W_t(l)L_t(j,l)dl + (1-\tau^K)R_t^k v_t(j)\bar{K}_{t-1}(j)\\
&\quad - \Psi(v_t)\bar{K}_{t-1}(j) + P_tZ_t(j) + \Pi_t(j)
\end{aligned}$$

where $W_t(l)$ is the nominal wage charged by the household for type l labor service. Consumption and labor income are, in nominal terms, subject to a sales tax $\tau^C > 0$ and a labor income tax $\tau^L > 0$, respectively.

Effective capital $K(j)$, which is subject to a rental income tax $\tau^K > 0$, is related to physical capital $\bar{K}(j)$ via $K_t(j) = v_t(j)\bar{K}_{t-1}(j)$, where $v_t(j)$ is the utilization rate of capital chosen by households and incurs a nominal cost of $\Psi(v_t)$ per unit of physical

capital.[13] Physical capital is accumulated by households according to $\bar{K}_t(j) = (1-\delta)\bar{K}_{t-1}(j) + u_t^i\left(1 - S\left(\frac{I_t(j)}{I_{t-1}(j)}\right)\right)I_t(j)$, where $0<\delta<1$ is the depreciation rate, $S(\cdot)I_t$ is an investment adjustment cost and u_t^i is an exogenous investment-specific efficiency shock.[14]

There are perfectly competitive labor packers that hire a continuum of differentiated labor inputs $L_t(l)$, pack them to produce an aggregate labor service and then sell it to intermediate goods producers. The labor packer uses the Dixit-Stiglitz aggregator for labor aggregation $L_t^d = \left(\int_0^1 L_t(l)^{1/(1+\eta_t^w)}dl\right)^{1+\eta_t^w}$, where L_t^d is the aggregate labor service demanded by intermediate goods producers, $L_t(l)$ is the lth type labor service supplied by all the households and demanded by the labor packer, and η_t^w is an exogenous wage markup shock.

For the optimal wage setting problem, we adopt the Calvo-pricing mechanism for nominal wage rigidities. Specifically, of all the types of labor services within each household, a fraction $0<\omega_w<1$ of wages cannot be changed each period. The wages for all other types of labor services follow a partial indexation rule $W_t(l) = W_{t-1}(l)\left(\pi_{t-1}e^{u_{t-1}^a}\right)^{\chi_w}(\pi e^{\gamma})^{1-\chi_w}$, where $W_{t-1}(l)$ is indexed by the geometrically weighted average of the growth rates of nominal wage in the past period and in the steady state, respectively. The weight $0<\chi_w<1$ controls the degree of partial indexation.

The household sector can be summarized by ten log-linearized equilibrium equations in terms of seventeen parameters $(h, \gamma, \alpha_G, \rho, \tau^C, \tau^K, \tau^L, \psi, \beta, \gamma, s, \delta, \xi, \omega_w, \chi_w, \rho_a, \eta^w)$, fifteen endogenous variables $(\hat{\lambda}_t, \hat{c}_t^*, \hat{c}_t, \hat{g}_t, \hat{R}_t, \hat{\pi}_t, \hat{P}_t^B, r_t^{\hat{k}}, \hat{v}_t, \hat{q}_t, \hat{i}_t, \hat{k}_t, \hat{\bar{k}}_t, \hat{w}_t, \hat{L}_t)$, and four exogenous shocks $(\hat{u_t^a}, \hat{u_t^b}, \hat{u_t^i}, \hat{u_t^w})$:

$$\begin{aligned}\text{Optimal consumption: } \hat{\lambda}_t = \hat{u_t^b} - \frac{h}{e^{\gamma}-h}\hat{u_t^a} - \frac{e^{\gamma}}{e^{\gamma}-h}\hat{c}_t^* \\ + \frac{h}{e^{\gamma}-h}\hat{c}_{t-1}^* - \frac{\tau^C}{1+\tau^C}\hat{\tau}_t^C\end{aligned} \tag{A.5}$$

$$\text{Composite consumption: } \hat{c_t^*} = \frac{c}{c+\alpha_G g}\hat{c}_t + \frac{\alpha_G g}{c+\alpha_G g}\hat{g}_t \tag{A.6}$$

$$\text{Consumption Euler: } \hat{\lambda}_t = \hat{R}_t + \mathbb{E}_t\hat{\lambda}_{t+1} - \mathbb{E}_t\hat{\pi}_{t+1} - \mathbb{E}_t\hat{u}_{t+1}^a \tag{A.7}$$

$$\text{Bond pricing: } \hat{R}_t + \hat{P}_t^B = \frac{\rho P^B}{1+\rho P^B}\mathbb{E}_t\hat{P}_{t+1}^B = \frac{\rho}{R}\mathbb{E}_t\hat{P}_{t+1}^B \tag{A.8}$$

$$\text{Optimal capital utilization: } r_t^{\hat{k}} - \frac{\tau^K}{1-\tau^K}\hat{\tau}_t^K = \frac{\psi}{1-\psi}\hat{v}_t \tag{A.9}$$

[13] Define the parameter $0<\psi<1$ such that $\frac{\Psi''(1)}{\Psi'(1)} \equiv \frac{\psi}{1-\psi}$.

[14] $S(\cdot)$ satisfies $S'(e^{\gamma}) = 0$ and $S''(e^{\gamma}) \equiv s > 0$.

$$\text{Optimal physical capital: } \hat{q}_t = \mathbb{E}_t\hat{\lambda}_{t+1} - \hat{\lambda}_t - \mathbb{E}_t\hat{u}^a_{t+1} + \beta e^{-\gamma}(1-\tau^K)r^k\mathbb{E}_t\hat{r}^k_{t+1} - \beta e^{-\gamma}\tau^K r^k\mathbb{E}_t\hat{\tau}^K_{t+1} + \beta e^{-\gamma}(1-\delta)\mathbb{E}_t\hat{q}_{t+1} \tag{A.10}$$

$$\text{Optimal investment: } \hat{i}_t = -\frac{1}{1+\beta}\hat{u}^a_t + \frac{1}{(1+\beta)se^{2\gamma}}\hat{q}_t + \hat{u}^i_t + \frac{\beta}{1+\beta}\mathbb{E}_t\hat{i}_{t+1} + \frac{\beta}{1+\beta}\mathbb{E}_t\hat{u}^a_{t+1} + \frac{1}{1+\beta}\hat{i}_{t-1} \tag{A.11}$$

$$\text{Effective capital: } \hat{k}_t = \hat{v}_t + \hat{\bar{k}}_{t-1} - \hat{u}^a_t \tag{A.12}$$

$$\text{Capital law of motion: } \hat{\bar{k}}_t = [1-(1-\delta)e^{-\gamma}]((1+\beta)se^{2\gamma}\hat{u}^i_t + \hat{i}_t) + (1-\delta)e^{-\gamma}(\hat{\bar{k}}_{t-1} - \hat{u}^a_t) \tag{A.13}$$

$$\text{Wage equation: } \hat{w}_t = -\kappa_w\left[\hat{w}_t - \xi\hat{L}_t - \hat{u}^b_t + \hat{\lambda}_t - \frac{\tau^L}{1-\tau^L}\hat{\tau}^L_t\right] + \frac{1}{1+\beta}\hat{w}_{t-1} + \frac{\beta}{1+\beta}\mathbb{E}_t\hat{w}_{t+1} + \frac{\chi_w}{1+\beta}\hat{\pi}_{t-1} - \frac{1+\beta\chi_w}{1+\beta}\hat{\pi}_t + \frac{\beta}{1+\beta}\mathbb{E}_t\hat{\pi}_{t+1} + \frac{\chi_w}{1+\beta}\hat{u}^a_{t-1} - \frac{1+\beta\chi_w-\rho_a\beta}{1+\beta}\hat{u}^a_t + \hat{u}^w_t \tag{A.14}$$

where $\kappa_w \equiv [(1-\beta\omega_w)(1-\omega_w)]/[\omega_w(1+\beta)(1+(1/\eta^w+1)\xi)]$, $\hat{\eta}^w_t \equiv \ln(1+\eta^w_t) - \ln(1+\eta^w)$, $\hat{\eta}^w_t$ is normalized to $\hat{u}^w_t \equiv \kappa_w\hat{\eta}^w_t$, $\hat{\tilde{u}}^i_t$ is normalized to $\hat{u}^i_t \equiv \frac{1}{(1+\beta)se^{2\gamma}}\hat{\tilde{u}}^i_t$, and λ_t is the Lagrange multiplier associated with the household's budget constraint. We set the capital, labor, and consumption tax rates to their constant steady states so that $\hat{\tau}^K_t = \hat{\tau}^L_t = \hat{\tau}^C_t = 0$.

Monetary and Fiscal Policy

The central bank implements monetary policy according to a Taylor-type interest rate rule. The government collects revenues from capital, labor, and consumption taxes, and sells nominal bond portfolios to finance its interest payments and expenditures. The fiscal choices must satisfy the government budget constraint $P^B_t B_t + \tau^K R^K_t K_t + \tau^L W_t L_t + \tau^C P_t C_t = (1+\rho P^B_t)B_{t-1} + P_t G_t + P_t Z_t$, where we have assumed the lump sum transfers are equal across households, i.e., $\int_0^1 Z_t(j)dj = Z_t$, and fiscal instruments follow the simple rules specified below.

The government sector can be summarized by seven log-linearized equilibrium equations in terms of thirteen parameters $(\tau^C, \tau^K, \tau^L, \beta, \gamma, \rho, \rho_r, \rho_g, \rho_z, \phi_\pi, \phi_y, \gamma_g, \gamma_z)$, sixteen endogenous variables $(\hat{b}_t, \hat{r}^k_t, \hat{k}_t, \hat{w}_t, \hat{L}_t, \hat{c}_t, \hat{\pi}_t, \hat{P}^B_t, \hat{g}_t, \hat{z}_t, \hat{y}_t, \hat{i}_t, \hat{v}_t, \hat{R}_t, \hat{s}^b_t, \hat{s}_t)$, and four exogenous shocks $(\hat{u}^a_t, \hat{u}^m_t, \hat{u}^g_t, \hat{u}^z_t)$:

$$\begin{aligned}\text{Government budget constraint:}\frac{b}{y}\hat{b_t} + \tau^K r^k \frac{k}{y}\left[\hat{\tau_t^K} + \hat{r_t^k} + \hat{k_t}\right] \\ + \tau^L w \frac{L}{y}\left[\hat{\tau_t^L} + \hat{w_t} + \hat{L_t}\right] \\ + \tau^C \frac{c}{y}(\hat{\tau_t^C} + \hat{c_t}) \\ = \frac{1}{\beta}\frac{b}{y}\left[\hat{b_{t-1}} - \hat{\pi_t} - \hat{P_{t-1}^B} - \hat{u_t^a}\right] \\ + \frac{b}{y}\frac{\rho}{\pi e^{\gamma}}\hat{P_t^B} + \frac{g}{y}\hat{g_t} + \frac{z}{y}\hat{z_t}\end{aligned} \tag{A.15}$$

$$\text{Aggregate resource constraint:}\hat{y_t} = \frac{c}{y}\hat{c_t} + \frac{i}{y}\hat{i_t} + \frac{g}{y}\hat{g_t} + \psi'(1)\frac{k}{y}\hat{v_t} \tag{A.16}$$

$$\text{Monetary policy rule:}\hat{R_t} = \rho_r \hat{R}_{t-1} + (1-\rho_r)\left(\phi_\pi \hat{\pi_t} + \phi_y \hat{y_t}\right) + \hat{u_t^m} \tag{A.17}$$

$$\text{Fiscal policy rule:}\hat{g_t} = \rho_g \hat{g}_{t-1} - (1-\rho_g)\gamma_g \hat{s_{t-1}^b} + \hat{u_t^g} \tag{A.18}$$

$$\hat{z_t} = \rho_z \hat{z}_{t-1} - (1-\rho_z)\gamma_z \hat{s_{t-1}^b} + \hat{u_t^z} \tag{A.19}$$

$$\begin{aligned}\text{Real primary surplus:}\hat{s_t} = \tau^K r^k \frac{k}{s}\left(\hat{\tau_t^K} + \hat{r_t^k} + \hat{k_t}\right) \\ + \tau^L w \frac{L}{s}\left(\hat{\tau_t^L} + \hat{w_t} + \hat{L_t}\right) \\ + \tau^C \frac{c}{s}(\hat{\tau_t^C} + \hat{c_t}) - \frac{g}{s}\hat{g_t} - \frac{z}{s}\hat{z_t}\end{aligned} \tag{A.20}$$

$$\text{Debt-to-output ratio:}\hat{s_t^b} = \hat{b_t} - \hat{y_t} \tag{A.21}$$

where $s_{t-1}^b \equiv \frac{P_{t-1}^B B_{t-1}}{P_{t-1} Y_{t-1}}$ denotes the market value of the debt-to-GDP ratio, $s = \tau^K r^k k + \tau^L w L + \tau^C c - g - z$, $0 < \rho_r, \rho_g, \rho_z < 1$ measure policy smoothness, $\phi_\pi, \phi_y > 0$ and γ_g, γ_z are policy parameters, and $(\hat{u_t^m}, \hat{u_t^g}, \hat{u_t^z})$ are exogenous policy shocks.

Exogenous Processes

All exogenous shocks follow autoregressive processes

$$\hat{u_t^s} = \rho_{es}\hat{u_{t-1}^s} + \epsilon_t^s, \qquad s \in \{a, b, i, p, w, m, g, z\} \tag{A.22}$$

where $\rho_{es} \in (0, 1)$ and the innovations ϵ_t^s are serially uncorrelated and independent of each other at all leads and lags.

Original System

Since the economy features a stochastic trend induced by the permanent technology shock A_t, some variables are not stationary. To induce stationarity, we therefore detrend these variables as: $y_t \equiv \frac{Y_t}{A_t}$, $c_t^* \equiv \frac{C_t^*}{A_t}$, $c_t \equiv \frac{C_t}{A_t}$, $k_t \equiv \frac{K_t}{A_t}$, $\bar{k_t} \equiv \frac{\bar{K_t}}{A_t}$, $i_t \equiv \frac{I_t}{A_t}$, $g_t \equiv \frac{G_t}{A_t}$, $z_t \equiv \frac{Z_t}{A_t}$, $b_t \equiv \frac{P_t^B B_t}{P_t A_t}$, $w_t \equiv \frac{W_t}{P_t A_t}$, $\lambda_t \equiv \Lambda_t A_t$. The model's equilibrium system in terms of the detrended variables can be summarized as follows.

Production function:

$$y_t \Delta_t^p = k_t^{\alpha} (L_t^d)^{1-\alpha} - \Omega \tag{A.23}$$

Capital-labor ratio:

$$\frac{k_t}{L_t^d} = \frac{w_t}{r_t^k} \frac{\alpha}{1-\alpha} \tag{A.24}$$

Real marginal cost:

$$mc_t = (1-\alpha)^{\alpha-1} \alpha^{-\alpha} (r_t^k)^{\alpha} w_t^{1-\alpha} \tag{A.25}$$

Intermediate goods producer's optimal price:

$$\begin{aligned} \mathbb{E}_t \Bigg[\sum_{k=0}^{\infty} (\beta \omega_p)^k \lambda_{t+k} \bar{y}_{t+k} \Bigg(\pi_t^* \prod_{s=1}^{k} \left(\frac{\pi_{t+s-1}}{\pi} \right)^{\chi_p} \frac{\pi}{\pi_{t+s}} \\ - (1 + \eta_{t+k}^p) mc_{t+k} \Bigg) \Bigg] = 0 \end{aligned} \tag{A.26}$$

Evolution of aggregate price index:

$$1 = (1-\omega_p)(\pi_t^*)^{-\frac{1}{\eta_t^p}} + \omega_p \left[\left(\frac{\pi_{t-1}}{\pi} \right)^{\chi_p} \frac{\pi}{\pi_t} \right]^{-\frac{1}{\eta_t^p}} \tag{A.27}$$

Optimal consumption:

$$\lambda_t (1 + \tau_t^C) = \frac{u_t^b}{c_t^* - \theta c_{t-1}^* e^{-u_t^a}} \tag{A.28}$$

Composite consumption:

$$c_t^* = c_t + \alpha_G g_t \tag{A.29}$$

Consumption Euler equation:

$$\lambda_t = \beta R_t \mathbb{E}_t \left[\frac{\lambda_{t+1} e^{-u_{t+1}^a}}{\pi_{t+1}} \right] \tag{A.30}$$

Bond pricing relation:

$$P_t^B = \mathbb{E}_t\left[\frac{1+\rho P_{t+1}^B}{R_t}\right] \tag{A.31}$$

Optimal capital utilization:

$$(1-\tau_t^K)r_t^k = \psi'(v_t) \tag{A.32}$$

Optimal physical capital:

$$q_t = \beta\mathbb{E}_t\left[\frac{\lambda_{t+1}}{\lambda_t}e^{-u_{t+1}^a}\left((1-\tau_{t+1}^K)r_{t+1}^k v_{t+1} - \psi(v_{t+1}) + (1-\delta)q_{t+1}\right)\right] \tag{A.33}$$

where q_t is the real price of capital in terms of consumption goods (i.e., Tobin's Q).

Optimal investment:

$$\begin{aligned} 1 =& q_t\tilde{u}_t^i\left(1 - S\left(\frac{i_t e^{u_t^a}}{i_{t-1}}\right) - S'\left(\frac{i_t e^{u_t^a}}{i_{t-1}}\right)\frac{i_t e^{u_t^a}}{i_{t-1}}\right) \\ &+ \beta\mathbb{E}_t\left[q_{t+1}\frac{\lambda_{t+1}e^{-u_{t+1}^a}}{\lambda_t}\tilde{u}_{t+1}^i S'\left(\frac{i_{t+1}e^{u_{t+1}^a}}{i_t}\right)\left(\frac{i_{t+1}e^{u_{t+1}^a}}{i_t}\right)^2\right] \end{aligned} \tag{A.34}$$

Effective capital:

$$k_t = v_t\bar{k}_{t-1}e^{-u_t^a} \tag{A.35}$$

Law of motion for capital:

$$\bar{k_t} = (1-\delta)e^{-u_t^a}\bar{k}_{t-1} + \tilde{u}_t^i\left(1 - S\left(\frac{i_t e^{u_t^a}}{i_{t-1}}\right)\right)i_t \tag{A.36}$$

Optimal wage:

$$\begin{aligned} &\mathbb{E}_t\left[\sum_{k=0}^{\infty}(\beta\omega_w)^k\lambda_{t+k}\bar{L}_{t+k}\left(w_t^*\prod_{s=1}^{k}\left(\frac{\pi_{t+s-1}e^{u_{t+s-1}^a}}{\pi e^{\gamma}}\right)^{\chi_w}\frac{\pi e^{\gamma}}{\pi_{t+s}e^{u_{t+s}^a}}\right.\right. \\ &\left.\left. - \frac{(1+\eta_{t+k}^w)u_{t+k}^b\bar{L}_{t+k}^{\xi}}{(1-\tau_{t+k}^L)\lambda_{t+k}}\right)\right] = 0 \end{aligned} \tag{A.37}$$

where

$$\bar{L}_{t+k} = \left[\frac{w_t^*}{w_{t+k}}\prod_{s=1}^{k}\left(\frac{\pi_{t+s-1}e^{u_{t+s-1}^a}}{\pi e^{\gamma}}\right)^{\chi_w}\frac{\pi e^{\gamma}}{\pi_{t+s}e^{u_{t+s}^a}}\right]^{-\frac{1+\eta_{t+k}^w}{\eta_{t+k}^w}}L_{t+k}^d \tag{A.38}$$

Evolution of aggregate wage index:

$$w_t^{-\frac{1}{\eta_t^w}} = (1-\omega_w)(w_t^*)^{-\frac{1}{\eta_t^w}} + \omega_w \left[\left(\frac{\pi_{t-1} e^{u_{t-1}^a}}{\pi e^{\gamma}} \right)^{\chi_w} \left(\frac{\pi e^{\gamma}}{\pi_t e^{u_t^a}} \right) w_{t-1} \right]^{-\frac{1}{\eta_t^w}} \tag{A.39}$$

Government budget constraint:

$$b_t + \tau_t^K r_t^k k_t + \tau_t^L w_t L_t + \tau_t^C c_t = \frac{1+\rho P_t^B}{P_{t-1}^B} \frac{b_{t-1}}{\pi_t e^{u_t^a}} + g_t + z_t \tag{A.40}$$

Aggregate resource constraint:

$$y_t = c_t + i_t + g_t + \psi(v_t) \bar{k}_{t-1} e^{-u_t^a} \tag{A.41}$$

Steady States

To solve for the steady states, we calibrate $\beta = 0.99$, $\alpha = 0.33$, $\delta = 0.025$, the average maturity of government bond portfolio $AD = 20$, $\eta^w = \eta^p = 0.14$, $g/y = 0.11$, $b/y = 1.47$, $\tau^C = 0.023$, $\tau^K = 0.218$, and $\tau^L = 0.186$. By assumption, $v = 1$, $\psi(v) = 0$, and $S(e^{\gamma}) = S'(e^{\gamma}) = 0$. The remaining steady states can be solved as follows.

From AD:

$$\rho = \left(1 - \frac{1}{AD}\right) \frac{1}{\beta} \tag{A.42}$$

From (A.30):

$$R = \frac{e^{\gamma}\pi}{\beta} \tag{A.43}$$

From (A.31):

$$P^B = \frac{\beta}{e^{\gamma}\pi - \rho\beta} \tag{A.44}$$

From $\tilde{u}^i = 1$ and (A.34):

$$q = 1 \tag{A.45}$$

From (A.33):

$$r^k = \frac{\frac{e^{\gamma}}{\beta} - (1-\delta)}{1-\tau^K} \tag{A.46}$$

From (A.32):

$$\psi'(1) = r^k(1 - \tau^K) \tag{A.47}$$

From (A.26):

$$mc = \frac{1}{1 + \eta^p} \tag{A.48}$$

From (A.25):

$$w = \left[mc(1 - \alpha)^{1-\alpha}\alpha^{\alpha}(r^k)^{-\alpha}\right]^{\frac{1}{1-\alpha}} \tag{A.49}$$

From (A.24):

$$\frac{k}{L} = \frac{\alpha}{1 - \alpha}\frac{w}{r^k} \tag{A.50}$$

From $\Delta^p = 1$, the final goods producer's zero profit condition, and (A.23):

$$\frac{\Omega}{L} = \left(\frac{k}{L}\right)^{\alpha} - r^k\frac{k}{L} - w \tag{A.51}$$

From (A.51):

$$\frac{y}{L} = \left(\frac{k}{L}\right)^{\alpha} - \frac{\Omega}{L} \tag{A.52}$$

From (A.35):

$$\bar{k} = ke^{\gamma} \tag{A.53}$$

From (A.36):

$$\frac{i}{L} = [1 - (1 - \delta)e^{-\gamma}]e^{\gamma}\frac{k}{L} \tag{A.54}$$

From (A.41):

$$\frac{c}{L} = \frac{y}{L}\left(1 - \frac{g}{y}\right) - \frac{i}{L} \tag{A.55}$$

From (A.40):

$$\frac{z}{L} = \left[\left(1 - \frac{R}{\pi e^{\gamma}}\right)\frac{b}{y} - \frac{g}{y}\right]\frac{y}{L} + \tau^C\frac{c}{L} + \tau^L w + \tau^K r^k\frac{k}{L} \tag{A.56}$$

From (A.29):

$$\frac{c^*}{L} = \frac{c}{L} + \alpha_G\frac{g}{y}\frac{y}{L} \tag{A.57}$$

From (A.37) and (A.39):

$$L = \left[\frac{w(1-\tau^L)}{(1+\tau^C)(1+\eta^w)}\frac{1}{(1-\theta e^{-\gamma})\frac{c^*}{L}}\right]^{\frac{1}{\xi+1}} \tag{A.58}$$

from which all level variables can be calculated from the steady state ratios given above.

Taking Model to Data

Define the private sector's one-step-ahead endogenous forecast errors as

$$\eta_t^x \equiv \hat{x}_t - \mathbb{E}_{t-1}\hat{x}_t, \qquad x \in \{\lambda, \pi, i, q, r^k, w, P^B\}. \tag{A.59}$$

The model consists of 36 log-linearized equilibrium equations and can be cast into the rational expectations system

$$\underbrace{\begin{bmatrix} \Gamma_0^{ee} & \Gamma_0^{ez} & \Gamma_0^{ed} \\ \mathbf{0} & I & \mathbf{0} \\ [\mathbf{0}, I] & \mathbf{0} & \mathbf{0} \end{bmatrix}}_{\underset{(36\times36)}{\Gamma_0(\theta^S)}} \underbrace{\begin{bmatrix} x_t^e \\ x_t^z \\ x_t^d \end{bmatrix}}_{\underset{(36\times1)}{x_t}} = \underbrace{\begin{bmatrix} \Gamma_1^{ee} & \Gamma_1^{ez} & \mathbf{0} \\ \mathbf{0} & P & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I \end{bmatrix}}_{\underset{(36\times36)}{\Gamma_1(\theta^S)}} \underbrace{\begin{bmatrix} x_{t-1}^e \\ x_{t-1}^z \\ x_{t-1}^d \end{bmatrix}}_{\underset{(36\times1)}{x_{t-1}}} + \underbrace{\begin{bmatrix} \mathbf{0} \\ I \\ \mathbf{0} \end{bmatrix}}_{\underset{(36\times8)}{\Psi}} \underset{(8\times1)}{\epsilon_t} + \underbrace{\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ I \end{bmatrix}}_{\underset{(36\times7)}{\Pi}} \underset{(7\times1)}{\eta_t} \tag{A.60}$$

where I denotes the identity matrix, $P = \mathrm{diag}\left(\rho_{ea}, \rho_{eb}, \rho_{ei}, \rho_{ep}, \rho_{ew}, \rho_{em}, \rho_{eg}, \rho_{ez}\right)$,

$$\underset{(21\times1)}{x_t^e} = [\hat{y}_t, \hat{c}_t^*, \hat{c}_t, \hat{k}_t, \hat{\bar{k}}_t, \hat{v}_t, \hat{L}_t, \widehat{mc}_t, \hat{b}_t, \hat{g}_t, \hat{z}_t, \hat{R}_t, \hat{s}_t^b, \hat{s}_t, \hat{\lambda}_t, \hat{\pi}_t, \hat{\imath}_t, \hat{q}_t, \hat{r}_t^k, \hat{w}_t, \hat{P}_t^B]$$

are the endogenous variables,

$$\underset{(8\times1)}{x_t^z} = [\hat{u}_t^a, \hat{u}_t^b, \hat{u}_t^i, \hat{u}_t^p, \hat{u}_t^w, \hat{u}_t^m, \hat{u}_t^g, \hat{u}_t^z]$$

are the exogenous shocks,

$$\underset{(7\times1)}{x_t^d} = [\mathbb{E}_t\hat{\lambda}_{t+1}, \mathbb{E}_t\hat{\pi}_{t+1}, \mathbb{E}_t\hat{\imath}_{t+1}, \mathbb{E}_t\hat{q}_{t+1}, \mathbb{E}_t\hat{r}_{t+1}^k, \mathbb{E}_t\hat{w}_{t+1}, \mathbb{E}_t\hat{P}_{t+1}^B]$$

are the conditional expectations of the last seven elements of x_{t+1}^e ,

$$\underset{(8\times1)}{\epsilon_t} = [\epsilon_t^a, \epsilon_t^b, \epsilon_t^i, \epsilon_t^p, \epsilon_t^w, \epsilon_t^m, \epsilon_t^g, \epsilon_t^z]'$$

are the shock innovations, and

$$\underset{(7\times1)}{\eta_t} = [\eta_t^\lambda, \eta_t^\pi, \eta_t^i, \eta_t^q, \eta_t^{r^k}, \eta_t^w, \eta_t^{P^B}]'$$

are the forecast errors.

Here the first row of (A.60) stacks the 21 structural equations (A.1)– (A.21), the second row stacks the 8 shock processes (A.22), and the third row stacks the 7 definitional equations (A.59). The unknown parameters θ consist of the structural parameters

$$\underset{(27\times 1)}{\theta^S} = [100\gamma, \xi, h, \alpha_G, \psi, s, \omega_p, \omega_w, \chi_p, \chi_w, \phi_\pi, \phi_y, \gamma_g, \gamma_z, \rho_r, \rho_g, \rho_z, \rho_{ea}, \rho_{eb}, \rho_{ei}, \rho_{ep}, \rho_{ew}, \rho_{em}, \rho_{eg}, \rho_{ez}, \bar{L}, \bar{\pi}]$$

and the volatility parameters

$$\underset{(24\times 1)}{\theta^V} = [\mu_a, \mu_b, \mu_i, \mu_p, \mu_w, \mu_m, \mu_g, \mu_z, \phi_a, \phi_b, \phi_i, \phi_p, \phi_w, \phi_m, \phi_g, \phi_z, \omega_a^2, \omega_b^2, \omega_i^2, \omega_p^2, \omega_w^2, \omega_m^2, \omega_g^2, \omega_z^2].$$

Conditional on θ^S and independent of the volatility processes, the above structural system can be solved by the procedure of Sims (2002) to deliver a linear solution of the form

$$\underset{(36\times 36)}{x_t} = \underset{(36\times 36)}{G(\theta^S)} x_{t-1} + \underset{(36\times 8)}{M(\theta^S)} \epsilon_t \tag{A.61}$$

which is then estimated over a vector y_t of 8 observable variables stacked in $y_{1:T} = [y_1, \ldots, y_T]'$, including log differences (denoted dl) of consumption, investment, real wage, government spending, and government debt; log (denoted l) hours worked, inflation, and nominal interest rate.[15] The observable variables are linked to the model variables x_t via the following measurement equations:

$$\underbrace{\begin{bmatrix} \text{dlCons}_t \\ \text{dlInv}_t \\ \text{dlWage}_t \\ \text{dlGovSpend}_t \\ \text{dlGovDebt}_t \\ \text{lHours}_t \\ \text{lInfl}_t \\ \text{lFedFunds}_t \end{bmatrix}}_{y_t} = \underbrace{\begin{bmatrix} 100\gamma \\ 100\gamma \\ 100\gamma \\ 100\gamma \\ 100\gamma \\ \bar{L} \\ \bar{\pi} \\ \bar{\pi} + 100(\gamma/\beta - 1) \end{bmatrix}}_{D} + \underbrace{\begin{bmatrix} \hat{c}_t - \hat{c}_{t-1} + \hat{u}_t^a \\ \hat{i}_t - \hat{i}_{t-1} + \hat{u}_t^a \\ \hat{w}_t - \hat{w}_{t-1} + \hat{u}_t^a \\ \hat{g}_t - \hat{g}_{t-1} + \hat{u}_t^a \\ \hat{b}_t - \hat{b}_{t-1} + \hat{u}_t^a \\ \hat{L}_t \\ \hat{\pi}_t \\ \hat{R}_t \end{bmatrix}}_{Zx_t}. \tag{A.62}$$

Let $\lambda_{1:T} = [\lambda_1, \ldots, \lambda_T]'$ contain all non-Gaussian latent states and $h_{1:T} = [h_1', \ldots, h_T']'$ contain all nonlinear latent states. In our empirical application ($T = 166$), $\lambda_{1:T}$ and $h_{1:T}$ have a total of 1,494 elements. In conjunction with the shock volatility specifications (2.2) and (2.3), equations (A.61) and (A.62) form a state space representation of the DSGE model whose conditional likelihood function $f(y_{1:T}|\theta, \lambda_{1:T}, h_{1:T})$ can be evaluated with the Kalman filter.

[15] See the Online Appendix of Leeper et al. (2017) for details on data construction.

Tables and Figures

- Table 5 lists the marginal prior distributions and the true values for the high-dimensional DSGE model under regime-M.
- Table 6 summarizes the posterior parameter estimates for the model of best fit.
- Table 7 reproduces the subsample posterior parameter estimates reported in Leeper et al. (2017), which are used to generate the simulated data sets in Sect. 5.3.
- Figures 8 and 9 display the autocorrelation function for each model parameter.
- Figures 10 and 11 compare the model's estimated parameters and volatilities with their true values, respectively.

Appendix B: Small-Scale DSGE Model

A log-linear approximation to the model's equilibrium conditions around the steady state can be summarized as follows:

$$\text{Dynamic IS equation:}\hat{c}_t = \mathbb{E}_t\hat{c}_{t+1} - \frac{1}{\tau}(\hat{R}_t - \mathbb{E}_t\hat{\pi}_{t+1} - \mathbb{E}_t\hat{z}_{t+1}) \tag{B.1}$$

$$\text{New Keynesian Phillips curve:}\hat{\pi}_t = \beta\mathbb{E}_t\hat{\pi}_{t+1} + \kappa\hat{c}_t \tag{B.2}$$

$$\text{Monetary policy:}\hat{R}_t = \rho_R\hat{R}_{t-1} + (1-\rho_R)[\psi_1\hat{\pi}_t + \psi_2(\hat{c}_t + \hat{g}_t)] + \epsilon_{R,t} \tag{B.3}$$

$$\text{Technology shock:}\hat{z}_t = \rho_z\hat{z}_{t-1} + \epsilon_{z,t} \tag{B.4}$$

$$\text{Government spending shock:}\hat{g}_t = \rho_g\hat{g}_{t-1} + \epsilon_{g,t} \tag{B.5}$$

Here $\tau > 0$ is the coefficient of relative risk aversion, $0 < \beta < 1$ is the discount factor, $\kappa > 0$ is the slope of the new Keynesian Phillips curve, $\psi_1 > 0$ and $\psi_2 > 0$ are the policy rate responsive coefficients, and $0 \leq \rho_R, \rho_z, \rho_g < 1$. Moreover, c_t is the detrended consumption, π_t is the inflation between periods $t-1$ and t, R_t is the nominal interest rate, z_t is an exogenous shock to the labor-augmenting technology that grows on average at the rate γ, and g_t is an exogenous government spending shock. Finally, the shock innovations $\epsilon_t = [\epsilon_{R,t}, \epsilon_{z,t}, \epsilon_{g,t}]'$ follow a multivariate Student-t distribution, i.e., $\epsilon_t \sim t_v(0, \Sigma_t)$, where $\Sigma_t = \text{diag}\left(e^{h_t}\right)$ and each element of $h_t = [h_t^R, h_t^z, h_t^g]'$ follows a stationary process

$$h_t^s = (1-\phi_s)\mu_s + \phi_s h_{t-1}^s + \eta_t^s, \qquad \eta_t^s \sim \mathbb{N}(0, \omega_s^2), \qquad s \in \{R, z, g\}. \tag{B.6}$$

The model is estimated over three observable variables, including log difference of consumption, log inflation, and log nominal interest rate. The observable variables are linked to the model variables via the following measurement equations,

Table 5 Priors and true values for high-dimensional DSGE model parameters

Structural parameters			Volatility parameters		
Name	Density (1, 2)	DGP	Name	Density (1, 2)	DGP
100γ	$\mathbb{N}$ (0.40, 0.05)	0.25	μ_a	t_ν (−7.0, 2.0)	−4.6
ξ	$\mathbb{G}$ (2.00, 0.50)	1.77	μ_b	t_ν (−7.0, 2.0)	−4.6
h	$\mathbb{B}$ (0.50, 0.20)	0.99	μ_i	t_ν (−7.0, 2.0)	−4.6
α_G	$\mathbb{U}$ (−1.75, 1.75)	−0.25	μ_p	t_ν (−7.0, 2.0)	−4.6
ψ	$\mathbb{B}$ (0.60, 0.15)	0.16	μ_w	t_ν (−7.0, 2.0)	−4.6
s	$\mathbb{N}$ (6.00, 1.50)	5.46	μ_m	t_ν (−7.0, 2.0)	−4.6
ω_p	$\mathbb{B}$ (0.50, 0.20)	0.92	μ_g	t_ν (−7.0, 2.0)	−4.6
ω_w	$\mathbb{B}$ (0.50, 0.20)	0.91	μ_z	t_ν (−7.0, 2.0)	−4.6
χ_p	$\mathbb{B}$ (0.50, 0.20)	0.06	ϕ_a	$\mathbb{B}$ (0.95, 0.01)	0.99
χ_w	$\mathbb{B}$ (0.50, 0.20)	0.18	ϕ_b	$\mathbb{B}$ (0.95, 0.01)	0.99
ϕ_π, regime-M	$\mathbb{N}$ (1.50, 0.20)	0.90	ϕ_i	$\mathbb{B}$ (0.95, 0.01)	0.99
ϕ_π, regime-F	$\mathbb{B}$ (0.50, 0.15)	n/a	ϕ_p	$\mathbb{B}$ (0.95, 0.01)	0.99
ϕ_y	$\mathbb{N}$ (0.125, 0.05)	0.10	ϕ_w	$\mathbb{B}$ (0.95, 0.01)	0.99
ρ_r	$\mathbb{B}$ (0.50, 0.20)	0.71	ϕ_m	$\mathbb{B}$ (0.95, 0.01)	0.99
γ_g, regime-M	$\mathbb{N}$ (0.15, 0.10)	0.26	ϕ_g	$\mathbb{B}$ (0.95, 0.01)	0.99
γ_z, regime-M	$\mathbb{N}$ (0.15, 0.10)	−0.11	ϕ_z	$\mathbb{B}$ (0.95, 0.01)	0.99
ρ_g	$\mathbb{B}$ (0.50, 0.20)	0.98	ω_a^2	$\mathbb{IG}$-2 (2.00, 0.05)	0.10
ρ_z, regime-F	$\mathbb{B}$ (0.50, 0.20)	n/a	ω_b^2	$\mathbb{IG}$-2 (2.00, 0.05)	0.10
ρ_{ea}	$\mathbb{B}$ (0.50, 0.20)	0.23	ω_i^2	$\mathbb{IG}$-2 (2.00, 0.05)	0.10
ρ_{eb}	$\mathbb{B}$ (0.50, 0.20)	0.40	ω_p^2	$\mathbb{IG}$-2 (2.00, 0.05)	0.10
ρ_{ei}	$\mathbb{B}$ (0.50, 0.20)	0.69	ω_w^2	$\mathbb{IG}$-2 (2.00, 0.05)	0.10
ρ_{ep}	$\mathbb{B}$ (0.50, 0.20)	0.74	ω_m^2	$\mathbb{IG}$-2 (2.00, 0.05)	0.10
ρ_{ew}	$\mathbb{B}$ (0.50, 0.20)	0.18	ω_g^2	$\mathbb{IG}$-2 (2.00, 0.05)	0.10
ρ_{em}	$\mathbb{B}$ (0.50, 0.15)	0.39	ω_z^2	$\mathbb{IG}$-2 (2.00, 0.05)	0.10
ρ_{eg}	$\mathbb{B}$ (0.50, 0.15)	0.13			
ρ_{ez}, regime-F	$\mathbb{B}$ (0.50, 0.15)	n/a	Fixed parameters		
$\bar{L}$	$\mathbb{N}$ (468, 5.00)	481.12	Regime-M: $\rho_z = 0.98$, $\rho_{ez} = 0.80$		
$\bar{\pi}$	$\mathbb{N}$ (0.75, 0.25)	0.60	Regime-F: $\gamma_g = \gamma_z = 0$		

Density (1, 2) refer to Gamma ($\mathbb{G}$), Normal ($\mathbb{N}$), Beta ($\mathbb{B}$), and Student-t_ν ($\nu = 2.1$ degrees of freedom) distributions with means and standard deviations indicated in parentheses; Uniform ($\mathbb{U}$) distribution with lower and upper bounds; Inverse-Gamma type-1 ($\mathbb{IG}$-1) distribution with parameters ν and s, where $p(\sigma) \propto \sigma^{-\nu-1} \exp\left(-\frac{\nu s^2}{2\sigma^2}\right)$; Inverse-Gamma type-2 ($\mathbb{IG}$-2) distribution with parameters α and β, where $p(\omega^2) \propto (\omega^2)^{-\alpha-1} \exp\left(-\frac{\beta}{\omega^2}\right)$. The effective priors are truncated at the boundary of the determinacy region

$$\begin{bmatrix} \text{dlCons}_t \\ \text{lInfl}_t \\ \text{lFedFunds}_t \end{bmatrix} = \begin{bmatrix} \gamma^{(Q)} \\ \pi^{(Q)} \\ \pi^{(Q)} + r^{(Q)} + \gamma^{(Q)} \end{bmatrix} + \begin{bmatrix} \hat{c}_t - \hat{c}_{t-1} + \hat{z}_t \\ \hat{\pi}_t \\ \hat{R}_t \end{bmatrix} \tag{B.7}$$

where $(\gamma^{(Q)}, \pi^{(Q)}, r^{(Q)})$ are connected to the model's steady states via $\gamma = 1 + \gamma^{(Q)}/100$, $\beta = 1/(1 + r^{(Q)}/100)$, and $\pi = 1 + \pi^{(Q)}/100$. Table 8 lists the marginal prior distributions for the small-scale DSGE model parameters.

Table 6 Posterior summary of DSGE model parameters

Structural parameters				Volatility parameters			
Name	Mean	90% HPD	Ineff	Name	Mean	90% HPD	Ineff
100γ	0.22	[0.16, 0.28]	4.4	μ_a	−0.33	[-0.74, 0.08]	4.9
ξ	1.65	[0.77, 2.41]	12.3	μ_b	6.41	[5.73, 7.06]	69.0
h	0.99	[0.98, 0.99]	103.6	μ_i	−0.93	[-1.56, -0.35]	5.8
α_G	−0.09	[-0.19, -0.00]	5.2	μ_p	−5.55	[-6.02, -5.04]	9.9
ψ	0.20	[0.11, 0.29]	9.0	μ_w	−3.63	[-4.27, -2.98]	13.7
s	7.35	[5.14, 9.57]	6.2	μ_m	−4.59	[-5.46, -3.76]	2.7
ω_p	0.91	[0.89, 0.94]	20.9	μ_g	0.79	[0.45, 1.18]	5.1
ω_w	0.84	[0.79, 0.89]	13.6	μ_z	1.78	[1.09, 2.46]	43.0
χ_p	0.09	[0.01, 0.16]	6.3	ϕ_a	0.95	[0.94, 0.97]	2.0
χ_w	0.07	[0.02, 0.11]	3.4	ϕ_b	0.95	[0.93, 0.96]	2.0
ϕ_π	1.00	[0.71, 1.31]	20.6	ϕ_i	0.95	[0.94, 0.97]	1.7
ϕ_y	0.17	[0.13, 0.22]	7.2	ϕ_p	0.95	[0.93, 0.97]	1.9
ρ_r	0.76	[0.70, 0.82]	15.7	ϕ_w	0.95	[0.93, 0.97]	1.7
γ_g	0.27	[0.16, 0.37]	17.8	ϕ_m	0.95	[0.94, 0.97]	2.0
γ_z	−0.07	[-0.20, 0.07]	39.9	ϕ_g	0.95	[0.93, 0.96]	2.3
ρ_g	0.98	[0.97, 0.99]	12.0	ϕ_z	0.95	[0.93, 0.97]	2.0
ρ_{ea}	0.24	[0.12, 0.35]	3.5	ω_a^2	0.02	[0.01, 0.04]	15.0
ρ_{eb}	0.38	[0.26, 0.50]	4.5	ω_b^2	0.02	[0.01, 0.04]	14.1
ρ_{ei}	0.64	[0.54, 0.73]	15.1	ω_i^2	0.05	[0.01, 0.10]	13.1
ρ_{ep}	0.74	[0.66, 0.80]	24.0	ω_p^2	0.03	[0.01, 0.06]	8.7
ρ_{ew}	0.45	[0.35, 0.58]	31.1	ω_w^2	0.05	[0.01, 0.08]	12.6
ρ_{em}	0.57	[0.46, 0.68]	16.1	ω_m^2	0.10	[0.03, 0.17]	7.5
ρ_{eg}	0.06	[0.01, 0.11]	3.9	ω_g^2	0.02	[0.01, 0.03]	24.9
$\bar{L}$	474.43	[473.31, 475.50]	23.9	ω_z^2	0.03	[0.01, 0.05]	19.2
$\bar{\pi}$	0.27	[0.00, 0.47]	3.4				

The posterior means and 90% highest probability density (HPD) intervals are computed using 10, 000 posterior draws from the TaRB-MH algorithm. The estimated model has regime-M in place and Student-t shocks with $\nu = 5$ degrees of freedom

Table 7 True values for DSGE model parameters

Name	Regime-M	Regime-F	Name	Regime-M	Regime-F
100γ	0.34	0.27	ρ_{ea}	0.23	0.30
ξ	1.78	2.25	ρ_{eb}	0.47	0.22
h	0.96	0.96	ρ_{ei}	0.76	0.47
α_G	−0.24	−0.38	ρ_{ep}	0.48	0.61
ψ	0.35	0.31	ρ_{ew}	0.36	0.20
s	7.08	3.47	ρ_{em}	0.52	0.87
ω_p	0.94	0.95	ρ_{eg}	0.16	0.29
ω_w	0.81	0.74	ρ_{ez}	0.80	0.90
χ_p	0.21	0.11	$100\sigma_a$	0.80	1.18
χ_w	0.19	0.06	$100\sigma_b$	14.54	16.95
ϕ_π	1.34	0.19	$100\sigma_i$	0.44	1.30
ϕ_y	0.16	0.21	$100\sigma_p$	0.09	0.13
ρ_r	0.79	0.37	$100\sigma_w$	0.27	0.23
γ_g	0.25	0	$100\sigma_m$	0.15	0.21
γ_z	−0.06	0	$100\sigma_g$	1.63	2.05
ρ_g	0.97	0.95	$100\sigma_z$	3.57	0.78
ρ_z	0.98	0.97			
$\bar{L}$	470.67	468.98			
$\bar{\pi}$	0.70	0.75			

The DGPs under regimes M and F correspond to their post-Volcker (1955:Q1–1979:Q4) and pre-Volcker (1982:Q1–2007:Q4) estimates reported in Leeper et al. (2017), respectively

Appendix C: Practical Guide

The subfolder `user/ltw17` contains the following files for the Leeper-Traum-Walker model, which are extensively annotated and can be modified, as needed, for alternative model specifications:

- `user_parvar.m`—defines the parameters, priors, variables, shock innovations, forecast errors, and observable variables.
- `user_mod.m`—defines the model and measurement equations.
- `user_ssp.m`—defines the steady state, implied, and/or fixed parameters.
- `user_svp.m`—defines the stochastic volatility parameters.
- `data.txt`—prepared in matrix form where each row corresponds to the observations for a given period.

Our MATLAB toolkit is readily deployed. Once the user supplies the above model and data files, the posterior distribution, marginal likelihood, and predictive distributions (and other quantities) are computed via the single function, `tarb.m`, as will be illustrated below. A printed summary of the results will be recorded in the MATLAB diary file `mylog.out`, which is saved to the subfolder `user/ltw17`.

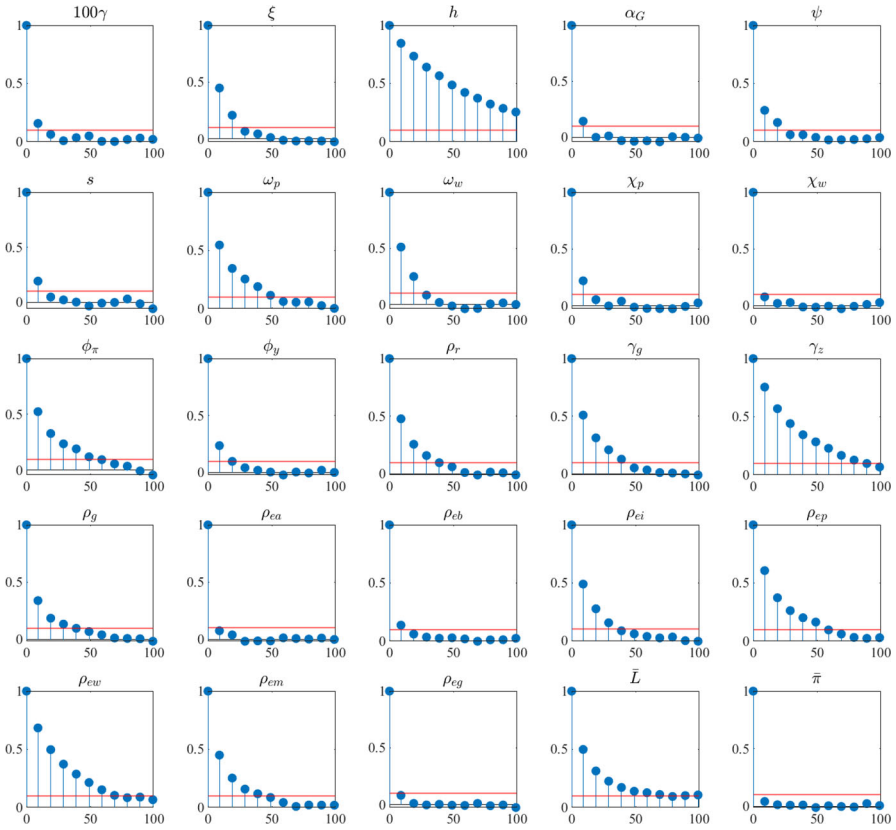


Fig. 8 Autocorrelation function of each structural parameter. *Notes*: Red horizontal lines correspond to an autocorrelation of 0.1

1. The prior sampling is implemented through the following block of code. The sampling results will be stored in the MATLAB data file `tarb_prior.mat`, which is saved to the subfolder `user/ltw17`.

```
npd = 10000;          % number of prior draws
T = 200;              % number of periods
sdof = 5;             % shock degrees of freedom
ltw17 = tarb(@tarb_spec,[],'npd',npd,'periods',T,'sdof',sdof);
tarb(@sample_prior,ltw17);
```

2. The posterior sampling is implemented in the following block of code. The estimation results will be stored in the MATLAB data file `tarb_full.mat`, which is saved to the subfolder `user/ltw17`.

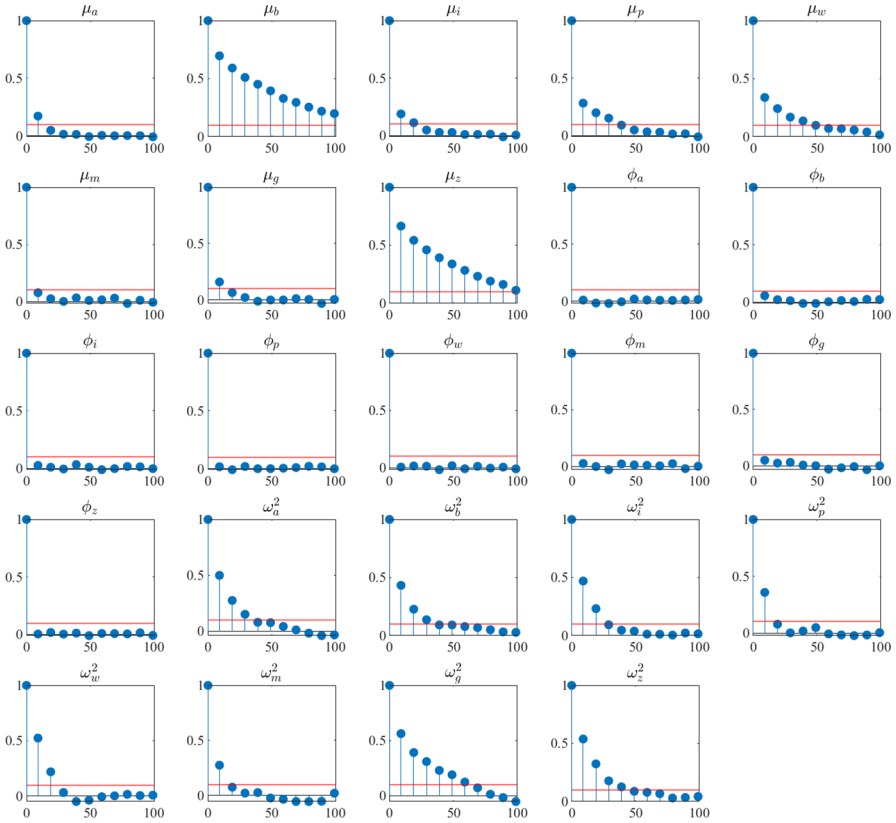


Fig. 9 Autocorrelation function of each volatility parameter. *Notes*: See Fig. 8

```
p = 0.7;              % blocking probability
w = 0.5;              % tailoring frequency
M = 11000;            % number of draws including burn-in
N = 1000;             % number of burn-in
ltw17 = tarb(@tarb_spec,ltw17,'prob',p,'freq',w,'draws',M,'burn',N);

tarb(@sample_post,ltw17);
```

3. The marginal likelihood estimation is implemented in the following block of code. The estimation results will be stored in the MATLAB data file `tarb_reduce.mat`, which is saved to the subfolder `user/ltw17`.

```
B = 7;                % number of blocks for structural parameters
ltw17 = tarb(@tarb_spec,ltw17,'blocks',B);
tarb(@marg_lik,ltw17)
```

4. Sampling the predictive distribution is specified in the following block of code.

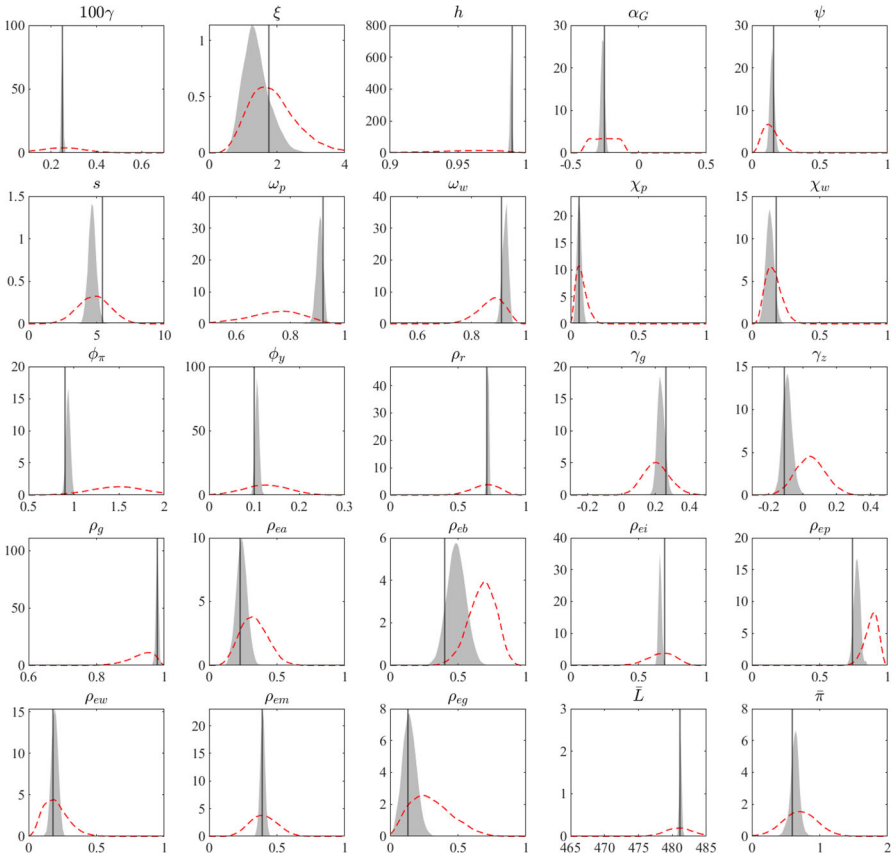


Fig. 10 Marginal prior and posterior distributions of each structural parameter. *Notes*: Each panel compares the prior (red dashed line) with the posterior (shaded area). Vertical lines denote the true parameter values. The kernel smoothed posterior densities are estimated using 10, 000 TaRB-MH draws

```
head = 50;          % training sample size
tail = 22;          % forecasting sample size
h = 1;              % forecasting horizon
ltw17 = tarb(@tarb_spec,ltw17,'datrange',[head tail],'pred',h);
```

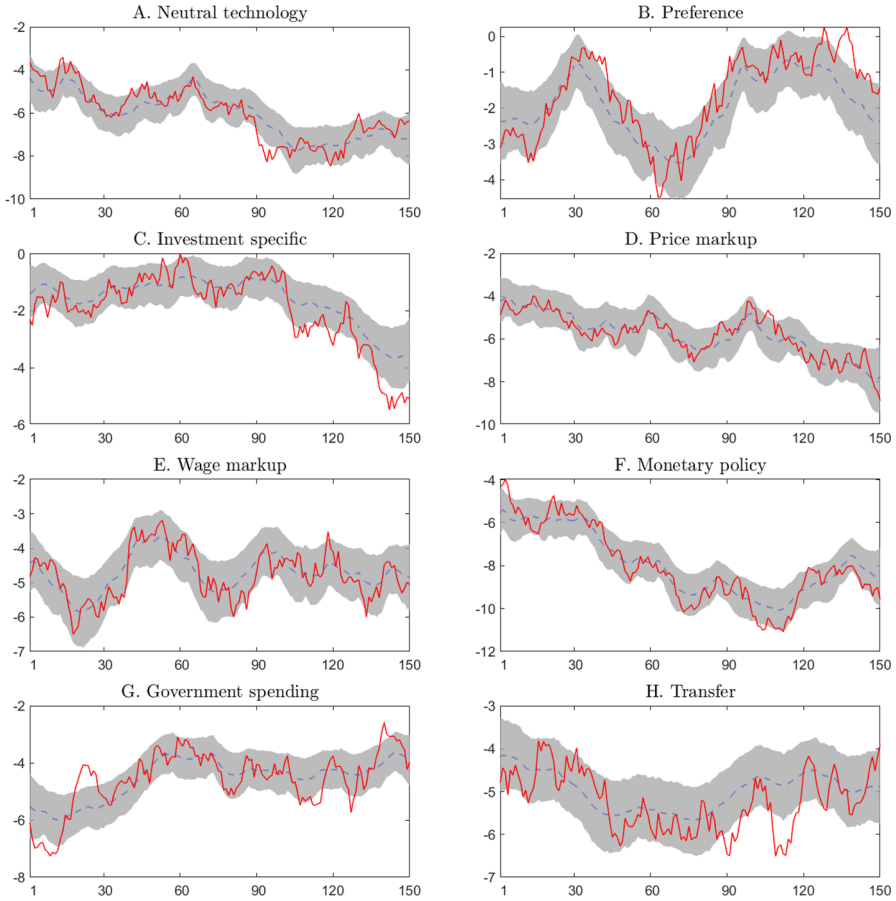


Fig. 11 Stochastic volatility of each shock innovation. *Notes*: Each panel compares the model's estimated log-variances with their true values (red solid line). Blue dashed lines denote median estimates, while shaded areas delineate 90% highest posterior density bands

Table 8 Priors for small-scale DSGE model parameters

Structural parameters		Volatility parameters	
Name	Density (1, 2)	Name	Density (1, 2)
τ	$\mathbb{G}\ (2.00, 0.50)$	μ_R	$t_\nu\ (-5.0,\ 2.0)$
κ	$\mathbb{G}\ (0.20, 0.10)$	μ_ζ	$t_\nu\ (-5.0,\ 2.0)$
ψ_1	$\mathbb{N}\ (1.50, 0.20)$	μ_g	$t_\nu\ (-5.0,\ 2.0)$
ψ_2	$\mathbb{G}\ (0.125, 0.05)$	ϕ_R	$\mathbb{B}\ (0.95, 0.01)$
$\gamma^{(Q)}$	$\mathbb{N}\ (0.40, 0.05)$	ϕ_ζ	$\mathbb{B}\ (0.95, 0.01)$
$\pi^{(Q)}$	$\mathbb{G}\ (0.75, 0.25)$	ϕ_g	$\mathbb{B}\ (0.95, 0.01)$
ρ_R	$\mathbb{B}\ (0.50, 0.20)$	ω_R^2	$\mathbb{IG}\text{-}2\ (2.00, 0.05)$
ρ_z	$\mathbb{B}\ (0.50, 0.20)$	ω_ζ^2	$\mathbb{IG}\text{-}2\ (2.00, 0.05)$
ρ_g	$\mathbb{B}\ (0.50, 0.20)$	ω_g^2	$\mathbb{IG}\text{-}2\ (2.00, 0.05)$

See Table 5

Funding F. Tan acknowledges the financial support from the Chaifetz School of Business summer research grant.

Code Availability The toolkit is publicly available at https://github.com/econdojo/dsge-svt.

Declarations

Conflicts of interest The authors declare that they have no conflict of interest.

References

Born, B., & Pfeifer, J. (2014). Policy risk and the business cycle. *Journal of Monetary Economics, 68*, 68–85.

Champagne, J., & Kurmann, A. (2013). The great increase in relative wage volatility in the United States. *Journal of Monetary Economics, 60*(2), 166–183.

Chan, J. C., & Grant, A. L. (2015). Pitfalls of estimating the marginal likelihood using the modified harmonic mean. *Economics Letters, 131*, 29–33.

Chen, R., & Liu, J. S. (2000). Mixture Kalman filters. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 62*(3), 493–508.

Chib, S. (1995). Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association, 90*, 1313–1321.

Chib, S., & Ergashev, B. (2009). Analysis of multifactor affine yield curve models. *Journal of the American Statistical Association, 104*(488), 1324–1337.

Chib, S., & Greenberg, E. (1995a). Hierarchical analysis of SUR models with extensions to correlated serial errors and time-varying parameter models. *Journal of Econometrics, 68*(2), 339–360.

Chib, S., & Greenberg, E. (1995b). Understanding the metropolis-hastings algorithm. *The American Statistician, 49*(4), 327–335.

Chib, S., & Jeliazkov, I. (2001). Marginal likelihood from the metropolis-hastings output. *Journal of the American Statistical Association, 96*(453), 270–281.

Chib, S., Nardari, F., & Shephard, N. (2002). Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics, 108*(2), 281–316.

Chib, S., & Ramamurthy, S. (2010). Tailored randomized block MCMC methods with application to DSGE models. *Journal of Econometrics, 155*(1), 19–38.

Chib, S., & Ramamurthy, S. (2014). DSGE models with student-t errors. *Econometric Reviews, 33*(1–4), 152–171.
Chib, S., & Zeng, X. (2020). Which Factors are risk factors in asset pricing? A model scan framework. *Journal of Business & Economic Statistics, 38*(4), 771–783.
Chiu, C.-W.J., Mumtaz, H., & Pinter, G. (2017). Forecasting with VAR models: Fat tails and stochastic volatility. *International Journal of Forecasting, 33*(4), 1124–1143.
Cúrdia, V., Del Negro, M., & Greenwald, D. L. (2014). Rare shocks, great recessions. *Journal of Applied Econometrics, 29*(7), 1031–1052.
Dave, C., & Malik, S. (2017). A tale of fat tails. *European Economic Review, 100*, 293–317.
Diebold, F. X., Schorfheide, F., & Shin, M. (2017). Real-time forecast evaluation of DSGE models with stochastic volatility. *Journal of Econometrics, 201*(2), 322–332.
Durbin, J., & Koopman, S. J. (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika, 89*(3), 603–615.
Franta, M. (2017). Rare shocks vs. non-linearities: What drives extreme events in the economy? Some empirical evidence. *Journal of Economic Dynamics & Control, 75*, 136–157.
Gelfand, A. E., & Dey, D. K. (1994). Bayesian model choice: Asymptotics and exact calculations. *Journal of the Royal Statistical Society Series B (Methodological), 56*(3), 501–514.
Geweke, J. F. (2005). *Contemporary Bayesian Econometrics and Statistics*. Wiley.
Herbst, E. P., & Schorfheide, F. (2016). *Bayesian Estimation of DSGE Models*. Princeton University Press.
Justiniano, A., & Primiceri, G. E. (2008). The time-varying volatility of macroeconomic fluctuations. *American Economic Review, 98*(3), 604–41.
Kapetanios, G., Masolo, R. M., Petrova, K., & Waldron, M. (2019). A time-varying parameter structural model of the UK economy. *Journal of Economic Dynamics & Control, 106*, 103705.
Kim, S., Shephard, N., & Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *The Review of Economic Studies, 65*(3), 361–393.
Kim, Y. M., & Kang, K. H. (2019). Likelihood inference for dynamic linear models with Markov switching parameters: On the efficiency of the Kim filter. *Econometric Reviews, 38*(10), 1109–1130.
Kulish, M., Morley, J., & Robinson, T. (2017). Estimating DSGE models with zero interest rate policy. *Journal of Monetary Economics, 88*, 35–49.
Lazarus, E., Lewis, D. J., Stock, J. H., & Watson, M. W. (2018). HAR inference: Recommendations for practice. *Journal of Business & Economic Statistics, 36*(4), 541–559.
Leeper, E. M., Traum, N., & Walker, T. B. (2017). Clearing up the fiscal multiplier morass. *American Economic Review, 107*(8), 2409–54.
Liu, X. (2019). On tail fatness of macroeconomic dynamics. *Journal of Macroeconomics, 62*, 103154.
Mele, A. (2020). Does school desegregation promote diverse interactions? An equilibrium model of segregation within schools. *American Economic Journal-Economic Policy, 12*(2), 228–257.
Rathke, A., Straumann, T., & Woitek, U. (2017). Overvalued: Swedish monteary policy in the 1930s. *International Economic Review, 58*(4), 1355–1369.
Sims, C. A. (2002). Solving linear rational expectations models. *Computational Economics, 20*(1), 1–20.
Sims, C. A., Waggoner, D. F., & Zha, T. (2008). Methods for inference in large multiple-equation Markov-switching models. *Journal of Econometrics 146*(2), 255–274. Nelson: Honoring the research contributions of Charles R.
Smets, F., & Wouters, R. (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review, 97*(3), 586–606.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Siddhartha Chib[1] · Minchul Shin[2] · Fei Tan[3,4]

✉ Fei Tan
tanf@slu.edu

[1] Olin Business School, Washington University in St. Louis, St. Louis, USA

[2] Research Department, Federal Reserve Bank of Philadelphia, Philadelphia, USA

[3] Department of Economics, Chaifetz School of Business, Saint Louis University, St. Louis, USA

[4] Center for Economic Behavior and Decision-Making, Zhejiang University of Finance and Economics, Hangzhou, China